

Budget blues for US science

President Bill Clinton's budget request for research falls short of the expectations of the scientific community, which have recently risen to unrealistic levels.

It has become customary for Al Gore, the US Vice-President and the strongest advocate for science and technology in the Clinton administration, to formally present the research portion of the president's annual budget proposal on the first Monday of February. This year not only Gore, but also Jack Lew, the director of the White House Office of Management and Budget, and Harold Varmus, the director of the National Institutes of Health (NIH), decided to give this presentation a miss.

By coincidence or not, the research budget presented on 1 February by Neal Lane, the president's science adviser, was perhaps less satisfying to the science lobby than any other of the Clinton era. Indeed, in his remarks, Lane came close to admitting his own disappointment with a set of proposals that would, if approved by Congress in their current form, actually see total research and development investment by the federal government fall during fiscal year 2000, which begins this October (see *Nature* **397**, 377; 1999).

The fall results from sharp reductions in spending on development and testing of military equipment; modest increases are proposed for civilian research, including an apparently generous 8 per cent rise for the National Science Foundation (NSF), which funds most non-biomedical research in US universities, and a much-decried 2 per cent increase for NIH. Important science programmes at other agencies, such as the Department of Energy (DoE) and NASA, generally get nothing more than the expected inflation rate of 2 per cent.

The figures themselves are not untypical of the budgets presented by Clinton during his first term. What has disappointed the science lobbyists this time is the fact that the national kitty is full (the overall budget now being in healthy surplus) and that rhetoric in favour of expanding the research base has been in full flood, not least from the president himself. During 1998, both houses of Congress joined with the administration in creating an apparent groundswell for such a decisive expansion.

This month's budget proposal shows just how hard such an expansion will be to sustain. Clinton himself thoroughly grasps the strategic importance of science to the country (as he grasps every issue he is

briefed on) but it is not a gut issue for him, in the manner of health care, adult literacy or education. He is not going to take a strong lead on the issue — and neither is Congress. Although various influential senators and congressmen are strongly attached to the science programmes that fall under their jurisdiction, the real leadership on the research issue on Capitol Hill over the past four years came from one Newt Gingrich, who resigned as Speaker of the House of Representatives last November.

It was Gingrich, for whom scientific research was one of the few legitimate tasks of the federal government, who decreed that appropriators should find the money to support biomedical research and other university research. This gave sympathetic committee chairmen the green light for last year's massive 15.6 per cent increase at NIH and for them to treat NSF's budget request more generously than in the past.

This year, the health-research lobby will obtain more for NIH from Congress (for which Clinton will shamelessly claim credit), but probably not the 15 per cent that it is seeking. The lobby's case is undermined by its overambitious objective of doubling the NIH budget to \$25 billion over the period 1998–2003 — an expansion that raises far too many unanswered questions. Does the talent exist to justify such a rapid expansion? Can NIH be relied upon to manage such an expansion properly? And what is the point of expanding at a rate that cannot be sustained in the longer term?

Given Congress's propensity to look after NIH, however, the community's first concern should be with the rest of the science budget. The NSF figure looks good, but most of it is for yet another new computing initiative; the NSF directorates, the workhorses of high-quality university science in the United States, remain more or less at level funding, as do the best programmes at NASA, DoE and elsewhere. If this administration were to follow through on its own science policy (see *Science in the National Interest*, OSTP, 1994) — dealing decisively with waste and politically motivated programmes in the federal government's oversized network of laboratories, and redirecting resources towards peer-reviewed, extramural basic research — these programmes could be sensibly expanded, at modest cost. □

What's up, postdoc?

A survey of junior researchers in Europe shows a moderately satisfactory situation, but room for improvement.

Postdoctoral workers, as any research supervisor knows, tend to be a stoical bunch. Overworked, undervalued and badly paid as they may be, the dominant picture that emerges from a survey of postdocs in Europe, commissioned jointly by *Nature* and the European Science Foundation, is that most seem prepared to accept their fate with little complaint (see page 640). As this survey reveals — and personal experience alone can testify — there are of course exceptions, both individually and geographically. But most seem relatively happy with their lot, accepting the drawbacks as, hopefully, a rite of passage.

Yet the survey also gives several causes for concern. One is the way that the experiences of postdocs vary widely across Europe, for example in the adequacy of research resources or levels of pay. This is not just a reflection of levels of economic development; it also reveals the

level of commitment that governments are prepared to make towards their researchers of the future. On this, the survey should give food for thought to those in the south of Europe in particular.

For those in the north, the issues are slightly different. Here, postdocs seem to be relatively well resourced and enjoy a good working environment — one reason for the south-to-north 'brain drain' in Europe that is itself a source of concern. But, as a separate news story indicates (see page 635), they are less prepared to accept the financial sacrifices that becoming a postgraduate student requires, particularly given the insecurity in their long-term career prospects. Here, the urgent need is to find ways of incorporating postgraduate students into a well-managed professional structure. Our survey reveals some of the issues that such efforts need to address. □



Debts and salaries tempt British graduates away from research

[LONDON] Increasing numbers of UK graduates are turning away from postgraduate training, according to new government data. Engineering and the physical sciences are worst hit, with financial reasons apparently at the root of the recruitment problems.

Over the past three years, the proportion of unfilled masters' and PhD studentships offered by Britain's research councils has more than doubled, from 4.8 to 11.0 per cent (see table). The Engineering and Physical Sciences Research Council (EPSRC) has the most empty places; this year 558 postgraduate training opportunities were unfilled.

Possible explanations include a healthy employment market, the low level of the PhD stipend, and the increasing debts of university graduates. EPSRC also points out that their figures do not include project studentships — PhDs offered as part of a research grant — whose numbers have been increasing over the last couple of years.

Money seems to be the key. "There is now a considerable difference between the stipend for research studentships and [graduate] starting salaries," comments David Schildt, Head of Cross Council Programmes at EPSRC.

Despite the recent £1,000 (US\$1,600) rise in PhD stipends, which came too late to influence the majority of this year's students,

Unfilled postgraduate positions

	1996-97		1997-98		1998-99	
	Offered	Unfilled	Offered	Unfilled	Offered	Unfilled
BBSRC	740	7	740	10	740	16
EPSRC	3,782	219	3,725	376	3,814	558
MRC	340	23	351	20	350	17
NERC	544	13	538	3	546	6
Total	5,406	262	5,354	409	5,450	597

BBSRC, Biotechnology and Biological Sciences Research Council; EPSRC, Engineering and Physical Sciences Research Council; MRC, Medical Research Council; NERC, Natural Environment Research Council. The 1998-99 figures for BBSRC, EPSRC and MRC are provisional.

the current minimum annual pay is still only £6,500. For PhD students, especially in engineering and physics, this is a fraction of their earning potential as graduates.

Swiftly rising levels of graduate debt — a new phenomenon in the United Kingdom — due to the gradual reduction in government grants to undergraduates make it more difficult for graduates to ignore lucrative job offers. This is likely to get worse; last year saw the introduction of £1,000 annual tuition fees.

The National Union of Students says that the mean debt of graduates is now between £4,000 and £5,000. This will worsen when the maintenance grant is removed at the end of 1999, when maintenance will be available only as loans. The union estimates that students starting university this year will leave with debts of £8,000–£12,000.

Another explanation is that the introduction of four-year undergraduate degrees with a strong research component is affecting the take-up of masters and PhDs.

The figures for unfilled studentships do not include figures for the Particle Physics and Astronomy Research Council, which is currently carrying over unfilled studentships from year to year because of the impact of four-year degrees on take-up.

Unease over the recruitment of research students was raised at a recent EPSRC conference. Richard Brook, chief executive of EPSRC, said these concerns were "very much shared", adding that, with the current level of the PhD stipend, research may appear increasingly less attractive.

"There has been more difficulty this year in filling the awarded studentships," admits Brook. Universities have been concerned about the issue of poor postgraduate recruitment for some time.

Some subjects have been harder hit than others. One speaker at the conference said that, despite a 25 per cent increase in applicants for computer science undergraduate degrees at the University of Southampton, it was "very hard" to attract postgraduates into academia, and even harder to keep them once they had completed PhDs.

A breakdown of EPSRC figures shows a rise in unfilled studentships across all subject areas. Schildt says that there are plans to address this by "increasing dialogue" between university departments, employers and students. "Every studentship place unfilled is a lost opportunity," he says.

The unfilled studentships take on added significance in the light of EPSRC's keenness to be judged on the knowledge — rather than the technical innovations — that it produces. Last week, departing chairman Alan Rudge denounced what he described as the "invention mania" behind pressure on universities to increase financial returns on inventions and intellectual property rights, which currently yield less than 1 per cent of university revenue.

Natasha Loder

See also the *News Analysis* on page 640.

Former Pasteur head defends politicians

[PARIS] The prosecution's case against former prime minister Laurent Fabius in France's 'contaminated blood' trial appeared to have been further undermined last week by testimony from François Gros, scientific adviser to Fabius in the 1980s.

Gros, one of France's most eminent biologists and a former director of the Institut Pasteur, absolved Fabius of any wrongdoing and assumed "full responsibility" for the actions underpinning the prosecution's case against the former prime minister.

Along with two other former ministers, Fabius has been charged with 'involuntary homicide' for alleged negligence in handling the risk of the transmission of HIV in the blood supply in 1985. The case against him rests mainly on the allegation that he delayed the introduction of a US test manufactured by Abbott Laboratories to protect the national market for a French test (see *Nature* 397, 548; 1999).

The central plank of this allegation is a decision taken at an interministerial meeting held in the prime minister's office

on 9 May 1985 to delay approval of the Abbott test "for some time". The prosecution brief states that "such a delaying measure could only be the translation of instructions given by the head of government".

Fabius, who introduced systematic screening of blood samples several months ahead of the United Kingdom and many other countries, has repeatedly denied that he delayed screening in the interests of protectionism. Testifying last week, Gros, who chaired the meeting, said that at no time was the prime minister involved in the discussions.

Gros also claimed that the aim of the meeting was to expedite introduction of screening, and that industrial considerations were "secondary". Gros, who himself faces trial for 'collusion in poisoning', over the affair, also contested the wider accusation that routine screening was delayed by protectionism, arguing that the speed of introduction depended on many factors, in particular costs to the healthcare system.

Declan Butler

US groups sue over approval of *Bt* crops...

[WASHINGTON] A coalition of US environmentalists and organic farmers last week filed a lawsuit against the US Environmental Protection Agency (EPA), accusing it of a "blatant disregard" for the law and its own regulations in its approval of the farming of genetically engineered crops.

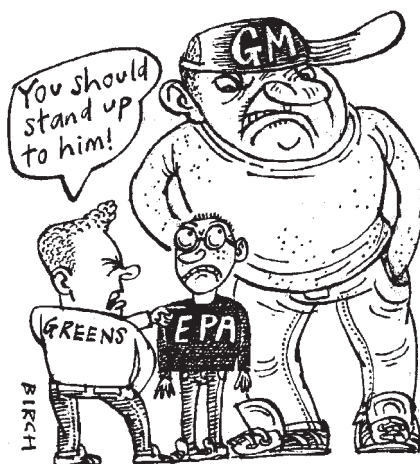
The 73 plaintiffs, led by Greenpeace, the Washington-based Center for Food Safety, and the International Federation of Organic Agricultural Movements, allege that since 1995 the EPA has approved *Bt* maize, cotton and potatoes — engineered to produce toxins made by the soil bacterium *Bacillus thuringiensis* — without fully assessing their environmental safety.

They contend that large-scale planting of the crops will cause insect resistance to the toxins, destroying their value for organic farmers, who have used them for almost forty years.

As a result, they argue, the crops present environmental hazards that should prevent EPA from approving them under federal pesticide law. The groups also claim that EPA disregarded the National Environmental Policy Act by failing to prepare an environmental impact statement analysing the effects of *Bt* crops.

The lawsuit, filed in the US District Court for the District of Columbia, makes the same arguments as the groups' 1997 petition demanding that EPA suspend existing registrations of *Bt* crops and stop issuing new ones until it has comprehensively studied the crops' environmental impact (see *Nature* 389, 317; 1997).

The plaintiffs say that the EPA has ignored the petition and that, with the fourth growing season of *Bt* crops now commencing, they must force the agency to take action. "The EPA is essentially studying this issue to death," says Joseph Mendelson, legal



director of the Center for Food Safety and lead counsel on the case.

In a written statement, the EPA said it is careful to ensure that the products it reviews are environmentally sound and beneficial, as required by law: "We believe the actions we've taken with regard to *Bt* will be sustained against this legal challenge."

Industry groups have been quick to attack the lawsuit. Jay Vroom, president of the American Crop Protection Association (ACPA), the major group representing US pesticide manufacturers — including most *Bt* seed producers — calls it "totally without merit" and argues that "EPA thoroughly assessed the safety of these products".

The group says that *Bt* crops are environmentally friendly: the toxins degrade rapidly and cannot bind in the digestive systems of non-target animals, and the crops spare farm workers exposure to chemical pesticides.

Since the EPA began registering *Bt* crops in 1995, it has approved seven varieties of maize, one of cotton and one of potatoes. The producers include Monsanto, Novartis Seeds, Mycogen, DeKalb and AgrEvo. In

1998, 20 million acres of *Bt* crops were planted in the United States, 15 million of which were maize, just under 20 per cent of US maize acreage.

The agency approves the crops as "plant pesticides" under the federal Fungicide, Insecticide and Rodenticide Act. The plaintiffs contend that EPA has broken this law in approving *Bt* crops, as an approved pesticide must not cause "unreasonable adverse effects" on the environment.

Although insect resistance to the crops has not yet been documented, the activists say they expect widespread resistance to evolve within a few years, because of the large acreages under cultivation and the selection pressure exerted by crops producing *Bt* constantly, as opposed to chemical spraying which was periodic. They also claim that the crops could hasten resistance by transferring the *Bt* trait to neighbouring plants, and argue that some *Bt* toxins may hurt beneficial insects and other non-target organisms.

In response, the ACPA says there is no scientific evidence for gene transfer to neighbouring plants, or for harm to non-target organisms. It adds that EPA requires resistance management plans. But these plans, which rely on areas of non-*Bt* crops to foster non-resistant, interbreeding insects, have been a prime target of activists. They claim they are inadequate to prevent resistance from emerging (see *Nature* 388, 817; 1997).

The activists' argument that EPA should have prepared an environmental impact statement for *Bt* crops faces an opposing legal precedent. A 1986 ruling stated that EPA approval of pesticides was not subject to the law requiring environmental impact statements. In that case, an Oregon man had sued to prevent EPA from registering seven herbicides that were sprayed along the road to his wife's farm.

Meredith Wadman

Spain makes transgenic crop producers pay into insurance fund

[BARCELONA] The Spanish government has decided that companies that produce or plant genetically modified crops must contribute to a 90 million Euro (US\$100 million) insurance fund intended to cover environmental accidents. The move reflects growing calls for tougher restrictions on such crops from opposition political parties, non-governmental organizations, and consumers' associations.

As a result of this pressure, the government's approach to transgenic crops will be debated in parliament this week after two left-wing parties expressed concern that Spain has authorized the planting of genetically modified crops that have not yet been approved in other countries of the European Community.

Environmental issues have become more controversial in Spain since last year's ecological disaster, when thousands of tons of toxic waste spilled into the Doñana national park last April after a retaining wall collapsed at the Aznalcóllar mines in Seville.

One party, the Bloque Nacionalista Gallego, is seeking either a moratorium or a strict limit on the import of such crops. The movement Ecologists in Action, which includes more than 300 environment-related organizations, has called for a ban on the 22 experimental field trials by the company Monsanto that have already been approved by the country's biosafety commission.

Concern has been triggered by the high importation of modified crops, especially

maize and soya. Between 15,000 and 20,000 hectares are said to have already been planted with such maize from the company Novartis. The number of licences for test plantings has increased from 36 in 1996 to 124 by January of this year.

Transgenic foodstuffs became an issue in Spain in 1996, when seven Greenpeace activists held a protest in Barcelona against a boat containing 45,000 tons of soya, 2 per cent of which was genetically modified.

Cristina Narbona, the environment-commission spokesperson of the socialist party PSOE, has urged the government to support demands being made at the biodiversity protocol meeting in Colombia to base the protocol on the so-called 'precautionary principle'.

Xavier Bosch

...as UK press reports come under fire

[LONDON] Leading British scientists are considering approaching the Press Complaints Commission over what they claim were inaccurate media reports on the risks of genetic modification. The move follows an unprecedented campaign in the British press during the past few weeks against genetically modified food.

The campaign has led to calls for the resignation of Britain's science minister, Lord David Sainsbury, as well as the suspension of the commercial planting of genetically modified crops until the risks to human health and the environment are better understood.

But, as the momentum of the past few weeks begins to subside, it has emerged that two national newspapers, *The Daily Telegraph* and the more populist *Daily Mail*, wrongly alleged that the government was suppressing reports on the risks of genetic modification.

In addition, *The Guardian* incorrectly reported that Sainsbury owns a patent on a virus promoter used in an experiment in which rats were alleged to have suffered from eating genetically modified potatoes. Scientists are now casting doubt on these findings, which also appeared in *The Guardian*.

The scientific community often disregards inaccurate media reports on science, partly to avoid discouraging journalists from writing about the subject. But many believe that a tougher approach is now needed to ensure that future reporting is more accurate and balanced.

"I have been very depressed by what I read over the past few weeks," says Ray Baker, chief executive of the Biotechnology and Biological Sciences Research Council. "We need to start talking to editors, and with the Press Complaints Commission."

These ideas are supported by senior officials at the Royal Society, as well as by John Beringer, professor of biological sciences at the University of Bristol and chairman of the government's Advisory Committee on Releases to the Environment (ACRE).

An ACRE report on the risks to farmland wildlife from modified crops was wrongly said to have been suppressed by the government in a front-page story in *The Daily Telegraph*. Beringer points out that ACRE reports are made public after each meeting.

Other concerned scientists include Nigel Poole, regulatory affairs manager at Zeneca Plant Science, who says he has been "stunned" by media hostility to the crops. Three years ago, he says, the press largely welcomed the development of a genetically modified tomato paste by Zeneca, cleared by the previous Conservative government. But the present Conservative opposition is one of the government's strongest critics.



Feeding frenzy: British newspapers have whipped up a storm over genetically modified crops.

Bridget Ogilvie, chair of the Committee on the Public Understanding of Science (COPUS), set up by the British Association, the Royal Institution and the Royal Society, believes that competition for readers may be one reason why Britain's national newspapers sometimes transform routine stories, even old news, into dramatic events.

For example, the *Daily Mail* last week used its front page to announce that a Royal Society report criticizing government policy on modified foods had been suppressed by the government. The society has distanced itself from the story, pointing out that the report has been in the public domain since September 1998 (see *Nature* **395**, 5; 1998).

"It is very difficult to know what to do,"

says Ogilvie. "We need an informed public debate on genetic modification. But how do you do that when newspapers are not interested?"

Over the past decade, COPUS has tried to strengthen links between scientists and the press, but these efforts have focused more on developing contacts with science correspondents, rather than the more influential political writers and senior editorial staff.

Contacts have also been mostly one-way: scientists have reached out to the press, but little attempt has been made to improve the quality of science reporting, particularly in the mass-circulation 'tabloid' newspapers. These sell more than 10 million copies a day, compared with the 3 million of the more serious 'broadsheets'.

The broadsheets employ specialist science correspondents who try to inject scientific perspective into their reports. But most of the coverage has come from political and environment correspondents, some of whom are openly critical of genetic modification.

Julie Hill, programme adviser of the Green Alliance, and a member of ACRE, says that some of the news coverage reflects public concerns. But she agrees that bad journalism should not go uncontested, although she thinks it is unrealistic to expect tabloid journalists to change their ways. Ehsan Masood

Britain reassures critics on risk research

[LONDON] The British government indicated last week that there will be no commercial planting of genetically modified crops until it is satisfied that the risk they pose to human health and the environment is negligible.

Ministers insist that this has always been the government's policy, but environmentalist pressure groups are claiming that it represents a climb-down in response to the largely media-generated pressure of the past two weeks for a three-year moratorium on commercial planting while the risks are properly assessed.

The government has created a sub-panel attached to its Advisory Committee on

Releases to the Environment to assess the risks of genetically modified crops to farm management practices and wildlife.

Last October, the government and the biotechnology industry agreed a one-year delay to the commercial introduction of herbicide-tolerant crops, and a three-year delay to insect-tolerant varieties, pending the completion of farm-scale trials.

The prospect of a further delay has been criticized by both industry and the scientific community, particularly as it could lead to companies and research councils cutting funding for plant biotechnology research.

Such a development, however, could result in

additional funds becoming available for research into biotechnology for healthcare, a view shared by Ray Baker, chief executive of the Biotechnology and Biological Sciences Research Council. "We'll have to put less money into [agricultural biotechnology] if there are no products at the end. We don't have the resources to do both."

Further delays could also boost the government's plans to promote biotechnology-based businesses, part of its support for more knowledge-based development. The Department of Trade and Industry is believed to want these businesses to focus on products in healthcare, rather than agriculture. E.M.

Japan on target to double science spend...

[TOKYO] Japan is likely to meet its goal of doubling its science spending by 2001, but it still needs to improve the research environment at national laboratories and universities. That is the conclusion of a progress report on the government's five-year plan on science and technology, written by the Science and Technology Agency and released last week.

The science and technology basic plan, which was launched in 1996, includes a commitment to increase the number of postdoctoral researchers to 10,000 over the course of five years. This target is now likely to be achieved by the end of the 1999 fiscal year, which starts in April.

But some other targets seem unlikely to be met. For example, a promise to increase

the number of non-Japanese researchers at national institutes is making slow progress: at present, there are only 0.4 such researchers per research group on average, less than half the original target.

The five-year plan was introduced to promote Japan's basic research and increase its science spending to a total of ¥17 trillion (US\$140 billion). With overall science spending expected to total ¥13.3 trillion during the 1999 fiscal year, the progress report says the financial target is likely to be met, provided that the budget for the fiscal year 2000 is kept at the current level (¥3.2 trillion), and that additional funding is provided through supplementary budgets.

Similarly, with the government creating

376 postdoctoral research posts, the number of postdocs in Japan is expected to total 10,187 in 1999, meeting the target a year earlier than planned.

The lack of foreign researchers is widely attributed to the general reluctance of research institutes to employ such researchers on a long-term basis. But many argue that the research environment in Japanese laboratories does not always appeal to overseas researchers.

"Unless these research institutes abandon the tenure system and employ a more flexible system allowing researchers to move easily from one research centre to another, their attraction will remain pretty low," says a British researcher working on a limited-term contract at one of the national institutes.

Another serious concern at universities and national research institutes is a shortage of laboratory technicians. Despite a commitment to provide one technician per researcher by 2001, the ratios of researchers to technicians at universities and national research institutes are currently 1:4 and 1:5, respectively. This means that young researchers, particularly postdocs, are often forced to act as lab technicians.

The report also says that little has been done to replace old buildings and facilities, particularly at national universities. It says a long-term strategy will have to be formed in the second phase of the plan, which begins in April 2001, to prevent the further deterioration of research facilities.

According to the report, 38 per cent of facilities at national research institutes, and 28 per cent of facilities at national universities, are in need of repair. At Kyoto University, for example, the science school has been waiting for a new building since 1995, even though it was listed as the top priority in the annual budget request.

"Manpower, location and funding are the three basic factors needed to carry out teaching and research at universities," says Kazuo Oike, dean of science at Kyoto University. "The one thing that's lagged behind is the location; the buildings have got old and dilapidated, and could present a potential danger to both students and staff."

The report says that the improvement of facilities will be one of the top priorities in the second phase of the plan, together with plans to increase the number of foreign researchers and to reinforce overall support for postdoctoral researchers.

But critics of the plan, especially in industry, say that too much emphasis has been placed on basic research at national universities and national research institutes, and that the second phase should focus more on application-orientated research to create new industries.

Asako Saegusa

... while Canada's budget boosts innovation

[MONTREAL] Canada is the latest country to declare its commitment to developing a 'knowledge-based economy'. Flush with a healthy fiscal surplus, finance minister Paul Martin last week presented a budget devoting Can\$1.8 billion (US\$1.2 billion) to encouraging research and innovation over the next three years.

The grand total of new spending is Can\$18 billion, and the share devoted to science and technology related items is second only to the health sector's. The health service will receive Can\$11.5 billion over the next five years, of which Can\$550 million over three years is new money for research. This includes plans for a coordinating body to be set up in 2000 and known as the Canadian Institutes for Health Research. The CIHR was first proposed last year by the Medical Research Council of Canada (see *Nature* 393, 613; 1998).

Four areas are identified as critical to the knowledge-based economy: creating knowledge, disseminating it, commercializing it, and hiring people to support it. One body to benefit from the new funding over the next three years will be the Canadian Space Agency, which will receive Can\$430 million on top of its previous budget, and Can\$300 million annually thereafter.

The Canada Foundation for Innovation will receive an extra \$200 million over this period. The foundation provides money for research infrastructure to hospitals, universities and research institutions.

Some Can\$240 million is allocated to setting up the CIHR, which will use networks to link all elements of the country's health-care system, providing a national focus for health research. The National Research Council, the three main federal research funding councils and the

National Health Research and Development Program will be given an extra \$150 million to back CIHR objectives.

Other extra spending includes Can\$60 million to improve researchers' access to computers, Can\$150 million to help companies market innovative products, and Can\$50 million for the Business Development Bank to finance knowledge-based and export-orientated companies.

An extra Can\$90 million will be allocated to the Networks of Centres of Excellence, which facilitate knowledge transfer among universities and the private sector.

As a result of the extra funding, the overall budget of the Medical Research Council (MRC) will grow in the next financial year to Can\$302.5 million, an increase of 11.4 per cent over this year. The 2000/2001 budget will be Can\$373.8 million, an increase over the current year of 37.7 per cent, and the 2001/2002 budget Can\$484.1 million, a massive 78.4 per cent increase over this year.

Reaction to the budget has been generally favourable. MRC president Henry Friesen hailed the creation of the CIHR as "historic". He said the government's largesse meant that the ever-widening gap between Canada and the United States "has taken a dramatically different course".

But critics said the overall increases for the health-care system would only bring its funding level back to where it had been three years previously, when the Liberal government started cost-cutting to eliminate the country's economic deficit.

Business people criticized the absence of corporate tax relief. Michael Wilson, executive director of the Fraser Institute, a Vancouver think-tank, said the lack of cuts to capital gains tax would have a serious impact on start-up companies.

David Spurgeon

Congress may block stem-cell research

[WASHINGTON] Seventy-seven anti-abortion members of Congress have written two letters to the Secretary of Health and Human Services, Donna Shalala, criticizing her for the recent decision to allow federal funding of research using human embryonic cells.

In a letter from members of the House of Representatives, 62 Republicans and eight Democrats call on Shalala to reverse the decision. They warn that if Harold Varmus, the director of the National Institutes of Health (NIH), proceeds as planned to fund the research, the NIH will be in violation of "both the letter and spirit" of the law banning federal funding for research in which human embryos are harmed or destroyed.

In a separate letter, seven Republican Senators accuse Shalala of a "unilateral attempt" to "effectively undermine congressional intent, by circumventing the current federal funding ban on embryo research".

Last month, Harriet Rabb, the general counsel of the Department of Health and Human Services (DHHS), issued a legal opinion that stem-cell research is exempt from the ban (see *Nature* **397**, 185; 1999). Rabb argued that because stem cells are not 'organisms' as defined in the law, federally funded scientists can work with them, even though the extraction of stem cells from embryos (which requires their destruction) cannot be publicly funded.

Varmus then announced that the NIH would begin funding stem-cell research in the coming months, under careful ethical supervision. But in a letter sent to Shalala on 11 February, the House lawmakers describe

Rabb's memo as "a carefully worded effort to justify transgressing" the ban on embryo research.

The letter's authors include the House majority leader, Dick Armey (Republican, Texas), the majority whip, Tom DeLay (Republican, Texas), and the chairman of the House Judiciary Committee, Henry Hyde (Republican, Texas).

Signatories to the letter sent by senators on 12 February include the majority whip, Don Nickles (Republican, Oklahoma). Laurie Boeder, a DHHS spokeswoman, says that Shalala will respond "in a timely fashion". She adds that Varmus has emphasized that, because of the potential benefits of the research for patients with a wide variety of diseases, "it's important for us to look into how we can appropriately fund federal research using existing stem cells".

But one biomedical lobbyist says "the political environment has changed dramatically". The letters "alert everybody that this group of people is going to fight very hard the ability to do stem-cell research under the existing statute".

The embryo research ban, first enacted in the 1996 fiscal year and renewed by Congress each year as part of annual spending bills that fund the NIH, bars federal support for "research in which a human embryo or embryos are destroyed".

The letter-writers claim that Rabb makes a specious distinction by reading the law narrowly to apply only to the act of destroying embryos, and not more broadly to include any research that depends on their destruc-

tion. "We prohibited the funding of research projects in which the lethal dissection or harmful manipulation of living human embryos is a necessary prerequisite," wrote the House members.

In testimony to the Senate in January, Varmus said he was unsure whether stem cells derived from the inner cell mass of the blastocyst might come together in culture to form an embryo. If so, the senators argue, then stem-cell research itself, regardless of issues surrounding their extraction, would be in direct violation of the embryo research ban.

Senator Arlen Specter (Republican, Pennsylvania), chair of the Senate subcommittee that funds the NIH and a strong supporter of stem-cell research, told the *New York Times* that the House letter puts proponents of the research "in very deep water".

Varmus plans to convene a subcommittee of his advisory committee to develop guidelines specifying what stem-cell work NIH can support. The subcommittee will also suggest a mechanism for an extra layer of review for stem-cell research proposals.

But this would become moot if Congress widens the ban to cover stem-cell research relying on the destruction of embryos. Jay Dickey (Republican, Arkansas), principal House author of the existing ban, issued a statement last week pointing out that the research relies on stem cells from embryos that were killed by having their stem cells removed. "This is precisely the kind of research for which we intended to ban, and did ban, federal funding." Meredith Wadman

Co-discoverer of evidence for quarks killed in diving accident

[BOSTON] Henry Kendall — professor of physics at the Massachusetts Institute of Technology (MIT), Nobel laureate and a tireless political activist — died last week while scuba diving in a Florida lake. He was 72 years old.

Kendall was an experienced deep-sea diver who had written books on the subject and designed underwater cameras. He had been photographing Wakulla Springs, the world's largest freshwater springs, accompanying a team of divers from the National Geographic Society.

Members of this team found him unconscious in shallow water a few feet from shore and took him to hospital, where he was pronounced dead on arrival. The cause of death is unknown, but his diving equipment is being tested to see if it malfunctioned.

Kendall received the Nobel prize for physics in 1990 with his MIT colleague Jerome Friedman and Richard Taylor of



Kendall: an activist as well as an academic.

Stanford University for work at Stanford in the late 1960s and early 1970s that provided the first direct evidence for quarks.

In electron-scattering experiments, they showed that protons and neutrons were 'lumpy', with point-like substructures in

their interior consistent with the quark model independently proposed in 1964 by Murray Gell-Mann and George Zweig. The finding paved the way for the Standard Model of physics.

Kendall will also be remembered for his activities on many political and environmental fronts. A co-founder of the Union of Concerned Scientists (UCS), and

its chairman since 1973, he was one of the first scientists to reveal safety flaws in the design and operation of nuclear power plants. He also warned against an unchecked nuclear arms race, and fought the 'Star Wars' defence initiative and space-based weapons.

Kendall co-authored books on energy policy, arms control, nuclear war and the fallacy of missile defence. More recently, he urged the adoption of measures to curb global warming — a subject on which he briefed President Bill Clinton in 1997.

"He always saw the big picture," says Friedman. Howard Ris, UCS executive director, says that Kendall "firmly believed that scientists could — and should — play an important role in public policy debates. His leadership ... was deeply rooted in the belief that, given accurate and credible information, the public and policymakers would, ultimately, make the right choices about the future."

Steve Nadis

Europe's young researchers seek proper rewards

Quirin Schiermeier

More than two thirds of Europe's young scientists say they are not given full credit for their research achievements. A survey commissioned jointly by *Nature* and the European Science Foundation reveals the countries with the best — and the worst — postdoc working conditions.

[LONDON] Britain and the Netherlands seem to be the favourite places for young researchers to work in Europe. More broadly, a survey carried out last year by e-mail reveals a clear north-south divide in the attractiveness of countries for postdoctoral and contract researchers.

But the survey also revealed complaints common to all parts of Europe. One is that young researchers do not feel they receive 'full credit' for their work. Overall, less than one third (31 per cent) say they felt fully credited, the number ranging from 17 per cent in Spain to 55 per cent in the Netherlands.

Asked about the reasons, most respondents cited hostile attitudes of colleagues, poor organization at their places of work, and excessive competitiveness.

Unsurprisingly, another complaint is that many researchers — particularly women — feel underpaid during their postdoctoral careers. One quarter (26 per cent) of men and more than one third (36 per cent) of women interviewed who have held more than one research contract say that payment is inadequate for their level of experience.

Only in the Netherlands (45 per cent) and Germany (38 per cent) do a reasonable proportion of young scientists consider their salaries to be appropriate. In other countries, financial dissatisfaction is the order of the day — particularly in Italy, where 64 per cent say they feel inadequately paid.

In the survey, 628 young researchers in physics, biology, chemistry and Earth sciences were asked about their current work, research goals, and the way their progress is monitored. The Strasbourg-based European Science Foundation's database of conference participants served as a sample.

Those surveyed were resident in eight European countries — the United Kingdom, Germany, France, Italy, the Netherlands, Sweden, Switzerland and Spain — and employed by universities, government funded research agencies, ministries or industry.

In general, young scientists say they enjoy considerable freedom and independence in choosing their research topics. More than 90 per cent of those polled — and a surprising 100 per cent in the Netherlands — consider their personal influence in setting research goals is sufficient or more than sufficient.

But the survey, which was carried out by IPSOS-RSL, also highlights what is troubling to many young researchers. It shows that there are considerable differences — by both country and sex — in the working conditions of young postdoctoral scientists in Europe. For example, more women (11 per cent) than men (5 per cent) find their personal influence on their research insufficient; one third of women (32 per cent), but almost half of the men (46 per cent), say it is more than sufficient.

Most of those surveyed reported that

their research goals are well defined, and that the extent to which their research progress is monitored is "about right". But in Italy, and among scientists of the lowest salary group (up to US\$16,000), almost one quarter say there is too little monitoring.

"The supervisor is very much hands off, and at times more critical assessment during the research would stop problems at a later date, for example when it comes to writing a publication," reports one scientist.

Some tend to think, in one researcher's words, that "nobody is better interested in what I am doing than me. But others feel disillusioned: "My boss basically does not care about things which are not directly related to his success." One respondent complains that "as a postdoc, I am simply a research tool, hired for a given period. My personal progress is irrelevant."

Young researchers in all the countries studied are required to carry out many tasks in addition to their scientific research. The most frequently named tasks are teaching, staff supervision, administration and preparing funding applications. Unsurprisingly, the demands to carry out additional tasks increase with increasing salary.

Although some find "the interaction between teaching and research very fruitful", most say their research suffers because they are distracted by non-research tasks. As one comments: "The breaking up of your research time into small bits due to other obligations leads to inefficient benchwork."

A French-based scientist points out that "teaching takes 292 hours a year in French universities. It is thus very difficult to have the time to build ambitious research projects, as well as to find the time to publish."

In Spain, many young scientists (42 per cent) complain that such tasks "interfere to a large extent" with research. In contrast, that opinion is shared by only 13 per cent in the Netherlands and 20 per cent in Britain.

Not all scientists feel they receive due credit for their achievements. "Being a postdoc is viewed as 'training'. All the credit for my research goes to my supervisors," says one.

Some see their fears coming true ("The old saying that women have to achieve at least twice as much as men to be given credit seems to be true"), while others are fatalistic: "It is what is expected of you, to be the work-horse and produce publications."

Young researchers in the Netherlands and in Britain, and male researchers overall, appear to have more freedom than their colleagues in Spain and Italy (and women) to explore new scientific ideas and question policies and arrangements that affect their own research.

One quarter (26 per cent) of those working in Italy, for example, say they have little or no freedom to do so, and only 17 per cent in Italy and Spain say they have complete freedom — compared with more than 30 per

Table 1 Opinion of employment package relative to level of experience among those who have held more than one research contract

	Country working in:						Working abroad		Total		
	UK	France	Germany	Spain	Italy	Netherlands					
Base	98	41	79	58	53	40		130		413	
Appropriate	24%	29%	38%	16%	8%	45%		29%		27%	
Fairly appropriate	40%	39%	34%	34%	21%	40%		42%		36%	
Inadequate	30%	22%	18%	40%	64%	13%		21%		29%	
No answer	6%	10%	10%	10%	8%	3%		8%		8%	
	Sex		Age			Gross salary (US\$000s)					
	Male	Female	25-29	30-34	35+	up to 16	16-23.9	24-31.9	32-39.9	40-47.9	48+
Appropriate	28%	25%	30%	27%	27%	8%	13%	25%	30%	33%	48%
Fairly appropriate	37%	33%	22%	34%	40%	19%	39%	35%	37%	40%	39%
Inadequate	26%	35%	33%	31%	27%	70%	39%	33%	25%	14%	11%
No answer	9%	6%	15%	8%	7%	3%	8%	8%	8%	12%	2%

cent in the Netherlands and in Britain. "Spain is the country with the strongest tradition against science," says one researcher.

"The system in Spain is very much boss-oriented," says an Englishman who works there. "It does not look well on initiatives from young people." Another scientist says: "I'm not supposed to have new ideas but rather just do what the group leader tells me to do. I'm not considered as a scientist but rather as a technical assistant."

Many postdocs also complain about inflexibility and egoistical attitudes of colleagues: "People are too selfish and everybody tries to stop the progress of others." Inflexible hierarchical structures, a fear of exploring new fields and the "tendency to stay on the fields with which people are comfortable" are often named as barriers to the exploration of new ideas.

A lack of personal independence and of permanent posts seems to be a problem, particularly outside the United Kingdom and the Netherlands. "A junior scientist is fully dependent on the favours and moods of established professors," says a German postdoc.

There are significant differences in how young scientists feel about the resources made available to them. Researchers in Germany seem able to rely on high-quality infrastructure, experimental equipment and technical staff. But this is not the case elsewhere. Technical staff are the cause of many complaints. In Spain and Italy, 70 per cent and 55 per cent of researchers, respectively, say that technical staff are "not very well" or "not at all well" provided for. One third of postdocs in both countries are also unhappy with the equipment available.

But what concerns young researchers most is a lack of time. More than one third (35 per cent) overall — and every second researcher in Spain and France — complain that they are not given enough time to complete research. Women claim to suffer significantly more from time pressure (45 per cent) than their male colleagues (32 per cent).

Similarly, the complaint is higher among chemists (45 per cent) than physicists (25 per cent), and — perhaps surprisingly — among researchers employed by government departments (47 per cent) than those paid by research funding agencies (31 per cent).

The most frequently cited reasons for lack of time are limited funding periods, inadequate technical support, equipment and software, and obligations additional to research, such as teaching, administration



Laboratory blues: many young scientists in Spain and Italy complain of inadequate technical support and poor equipment, while those who work in Germany enjoy the best resources available in Europe.

and preparing grant applications.

Women in particular find that family duties can interfere with scientific ambition. "Because of my child, I work only eight hours a day, which is quite unusual in science," says a female postdoc. "Having a family makes you less competitive."

Some blame the "bad habit" of conducting several research projects in parallel. "I think we spend too much time planning and assessing our work," says one Swedish researcher. "We should spend some of that time actually performing the work."

One postdoc comes to the sobering conclusion that "my ambitions are greater than my skills". Many young scientists suffer from the fact that postdoctoral research does not seem to offer a clear career perspective.

Comments included: "Realizing that doing good research will not greatly increase future job opportunities is demotivating"; and "Uncertainty about the future interferes with the ability to concentrate on research".

Other typical problems encountered in research are a lack of discussion of ideas with other groups, low motivation of graduate students involved in the research, and publication pressure ("just the quantity counts"). There is also the "politicization" of research:

"Prestige is more and more [important]," says one. "You may end up [appearing to be] a successful scientist, but you are just a successful politician who has the right contacts."

As for mobility, one third of those surveyed work outside their country of origin. About two thirds say this does not cause difficulties, with women being more adaptable than men. Where problems arise, they are mostly related to such issues as language, culture, work permits, taxes and pensions.

Unsurprisingly, the satisfaction of young researchers strongly depends on the working conditions at their host institutes. The fortunate ones enthuse about their freedom to carry out research, about the support and the credit they are given, and about the academic lifestyle that allows them to interact with like-minded people from many countries. The difficulties of the less fortunate revolve around such issues as career uncertainty, hierarchical scientific structures, and inadequate salaries, funds and equipment.

If several problems coincide, frustration increases rapidly. "I will leave for the United States if in a reasonable time I do not get a permanent position, or if I get fed up with the attempts of some boss to [add his or her name to] publications without making any contribution," says one postdoc.

The United Kingdom, where almost 50 per cent of young researchers come from abroad, is by far the most attractive European host country, in absolute and relative terms. In Spain and Italy, in contrast, only one in ten young researchers is of foreign origin — a fact that may reflect some of the other factors revealed by this survey. □

Table 2 Extent to which credit received for research achievements

	UK	France	Country working in:				Working abroad	Total
			Germany	Spain	Italy	Netherlands		
Base	132	74	106	83	93	56	198	628
Full credit	38%	27%	35%	17%	26%	55%	32%	31%
Reasonable credit	48%	64%	51%	61%	53%	34%	59%	53%
Little credit	12%	8%	12%	14%	19%	11%	8%	13%
No credit	-	-	1%	1%	-	-	-	-

Tax perks for research should be permanent, say US congressmen

[WASHINGTON] The leaders of the Science Committee in the US House of Representatives have introduced legislation that would permanently extend the research and development tax incentive, a mechanism for encouraging the expansion of industrial R&D. This is currently renewed by Congress on an annual basis.

Jim Sensenbrenner (Republican, Wisconsin), chairman of the committee, and George Brown (Democrat, California), its senior minority member, have co-sponsored the legislation, arguing that the tax incentive is a "tried and true" success that would be even more effective if corporations could rely on its permanence. Senator Pete Domenici (Republican, New Mexico) and Jeff Bingaman (Democrat, New Mexico) are planning to introduce similar legislation in the Senate.

The R&D tax credit is strongly supported by large US industrial corporations. Sceptics say this provides a strong disincentive for Congress to make it permanent; if it is merely renewed on an annual basis, members of the House Ways and Means and Senate Finance committees can plead for

contributions each year from such corporations in exchange for supporting renewal.

Japanese dioxins scare leads to health probe

[TOKYO] Keizo Obuchi, Japan's prime minister, is to set up a cabinet-level working group to study the effects of dioxins on health, focusing on potential carcinogenic properties. The move follows a television report claiming that high concentrations of dioxins were detected in vegetables grown in Saitama prefecture, near Tokyo.

The report caused the prices of vegetables produced in Saitama to plummet, but it was later discovered that dioxins were detected only in tea leaves. Obuchi, who ate Saitama-grown vegetables on television to demonstrate their safety, said at a press conference that the panel will try to establish a national standard for a safe and acceptable level of dioxins.

French embryo report fails to rule on ban

[PARIS] A report soon to be released by the Parliamentary Office of Evaluation of Scientific and Technological Choices leaves open the question of whether the French ban

on research using human embryos should be lifted.

The issue is likely to be at the centre of parliamentary debate later this year, when France is scheduled to revise the battery of bioethics legislation it adopted in 1994. The parliamentary report, based on hearings of more than 60 scientists, physicians, legal experts and bioethicists, is expected to stop short of taking a position on the question, instead arguing that it should merely set out the arguments for and against lifting the ban.

Campbell quits Britain's genetics advisory panel

[LONDON] Sir Colin Campbell, vice-chancellor of the University of Nottingham, has stepped down as chairman of the British government's Human Genetics Advisory Commission (HGAC), following his decision to take up a senior non-executive post with a major insurance company.

A statement from the commission says that Campbell has left to "remove himself from any possible future embarrassment" in the event of an HGAC investigation into genetics and insurance. Onora O'Neill, principal of Newnham College, University of Cambridge, and chair of the trustees of the Nuffield Foundation, will be acting chair of the HGAC until June.

Germany enters the supercomputer age

[MUNICH] Germany's first supercomputer, able to perform one billion calculations per second, will be installed later this year at the Bavarian Academy of Science in Munich. The DM90 million (US\$50 million) high-performance computer will be among the ten most powerful systems worldwide.

The new facility will be accessible to the entire scientific community in Germany. According to Wolf-Dietrich Schubring, head of the academy's computing department, the most urgent need for computer simulation of complex natural processes lies in such areas as climatology, hydrodynamics, materials sciences, biomedicine, combinatorial chemistry, solid-state physics and astrophysics.

Call for more research on endocrine disruptors

[WASHINGTON] **Although no direct evidence links declines in male sperm counts to so-called 'endocrine disruptor' chemicals in the environment, concern over this and other health effects is still warranted, according to a survey of research published in the journal *Pure and Applied Chemistry*.**

More studies, and better methods of

screening and testing environmental contaminants, are needed to settle questions about the safety of environmental oestrogens, according to the report. The year-long survey involved more than two dozen authors and was sponsored by the International Union of Pure and Applied Chemistry and the International Unions of Pharmacology and of Toxicology, with the support of the International Council for Science.

US cash to aid Soviet scientists missed target

[WASHINGTON] The US General Accounting Office (GAO) has issued a report highly critical of two Department of Energy programmes intended to help experts on nuclear, biological and chemical warfare in the former Soviet Union move into civilian research. Most of the money spent on the five-year-old Initiative for Proliferation Prevention and the newer Nuclear Cities Initiative went into the department's own laboratories in the United States; not all of the money sent to the former Soviet Union actually reached scientists, and there is "some evidence" that scientists were still working on weapons projects.

The GAO was asked by Senator Jesse Helms (Republican, North Carolina) to

review both programmes. Senator Helms is a long-standing critic of US efforts to offer technical or other assistance to the former Soviet states.

School students could pilot their own satellite

[MUNICH] The German-led X-ray satellite



ROSAT could take on a new role next year as a teaching aid, following the decision to abandon its scientific activities after eight successful years of collecting scientific data. The idea

has been suggested by Joachim Trümper, director of the ROSAT project and of the Max Planck Institute for Extraterrestrial Physics.

Earlier this month, attempts were abandoned to reactivate the German-developed satellite for a final observation campaign, which was initially expected to collect data for only 20 months, after a number of faults developed during the previous observation period (see *Nature* 395, 826; 1998). Trümper has proposed to the German Space Agency that the satellite should be used as a 'school satellite', steerable by pupils and laymen, during next year's world exhibition EXPO 2000 in Hanover.

Trials and tribulations of PhD students

Sir — Your reports of the suicide of Jason Altom, a student at Harvard University, crystallized some thoughts about the frustrations of many graduate students (*Nature* **395**, 823, 826; 1998 & **397**, 291; 1999).

Tragedies like this are known at almost all research universities. It is surprising how much control supervisors can exert over their students' careers, and bad relations with supervisors are nightmarish. There are many factors that individually may seem innocent, but which can have disastrous cumulative effects. One of these is the system of recommendations (references).

Throughout a postgraduate scientific career, recommendations are required, and that of one's PhD supervisor is almost indispensable. Applicants will obviously try to submit only positive recommendations, so this system is inherently biased and subjective. As long as one has a friendly and objective supervisor, all is well. But a student who has strained relations with their boss is at an unfair disadvantage.

I know of many students who cannot speak out against their mentors even after

finishing their PhD, largely because they will need references in the future. This breakdown might be read as an inability to maintain professional relations, but such relations are a two-way street.

References should be less important, and hiring choices should be decided by more objective parameters such as test scores, quality of work and publications, and personal interviews. Sympathetic professors and graduate programmes should try to put this into practice as soon as possible, and encourage others to do so.

Sudhanshu Dole

*Institute for Genetics, University of Cologne,
121 Weyertal, Cologne 50931, Germany*

Sir — When I was an MSc student in Canada, I was the president of 100 graduate students, and frequently heard about or counselled those who had suffered verbal abuse, threatening remarks, and many other abuses at the hands of their supervisors. The problem is a Pandora's box.

After hearing so many complaints, I devised a form on which graduate students could evaluate their supervisors. But it was

overwhelmingly turned down by the committee of my fellow students, who were afraid that negative remarks would harm their relationships with their supervisors and subsequent letters of reference. At that point, I became known among my fellow students as the guy who wanted to expose departmental practices.

I then moved to a prominent university in England to study for a PhD. I soon learned that things can be even worse. There is rarely any undergraduate evaluation here, and the role of graduate supervisors (of which there is almost always only one) is usually minimal. The supervisor gives you a project and expects results — plain and simple. Fortunately, my supervisor cares about my development as a scientist, but he is a rare example.

I believe that further cases like that of Altom can be avoided if prospective students know the track record of the proposed supervisor and the potential for abuse. I urge the formation of an international, independent directory of graduate evaluations of supervisors.

Name and address supplied

How hunger keeps the population down

Sir — Stephen G. Warren and Roger Short leave the reader puzzled about how the volume of agricultural output sets limits to the size and growth of human population (*Nature* **397**, 101; 1999). Both authors agree with Malthus that such limit-setting occurs, but Short shows that it does not operate through hunger reducing women's fertility. But surely famine and malnutrition sharply increase child mortality, thereby reducing recruitment to the ranks of sexually mature women and so rendering population growth impossible.

The increase in food production began in Europe in the seventeenth and eighteenth centuries and, at least initially, owed more to better organization than to science.

Hermann Bondi

Churchill College, Cambridge CB3 0DS, UK

Revolutionary ideas come round again

Sir — *Plus ça change*. Editorials in *Nature* pointing out that the studies of many molecular biologists have been largely qualitative are one thing, and very welcome.

However, it is quite wrong to comment³ that "the biologist ... will probably have had very little quantitative training", since at least those biochemists who deal with enzymes and metabolism must of necessity work properly with numbers.

The post-genomic era will certainly be characterized by large-scale and quantitative analyses, but it is the large scale that is new, not the quantitative aspect. Sensitivity analysis, which is what Alon *et al.*⁴ have done, has been a tool of biochemistry for a quarter of a century⁵⁻⁷. Robustness and rigidity against large changes in fluxes following changes in enzyme levels are a well-established property of metabolic networks⁸ that follows naturally from their structure and kinetic properties⁹.

But etymologically you are correct; if quantitative methods in biology are a "revolution" it is only because they are coming round again.

Douglas B. Kell

*Institute of Biological Sciences, Cledwyn Building,
University of Wales, Aberystwyth SY23 3DD, UK*

1. Maddox, J. *Nature* **355**, 201 (1992).

2. Maddox, J. *Nature* **368**, 95 (1994).

3. *Nature* **397**, 89 (1999).

4. Alon, U. *et al.* *Nature* **397**, 168–171 (1999).

5. Kacser, H. & Burns, J. A. *Symp. Soc. Exp. Biol.* **27**, 65–104 (1973).

6. Heinrich, R. & Rapoport, T. A. *Eur. J. Biochem.* **42**, 89–95 (1974).

7. http://gepasi.dbs.aber.ac.uk/metab/mca_home.htm

8. Stephanopoulos, G. & Vallino, J. J. *Science* **252**, 1675–1681 (1991).

9. Mendes, P. *et al.* *Biochim. Biophys. Acta* **1289**, 175–186 (1996).

Patent attorneys pay their way

Sir — A journal of *Nature's* calibre should be above taking cheap shots at attorneys in general and patent attorneys in particular. In your editorial on the sharing of research materials, advocating the use of uniform materials transfer agreements, the final sentence was unnecessary, unprofessional, and sorely weakened an otherwise cogent, if misguided, argument (*Nature* **396**, 97; 1998).

"No one need lose but the lawyers"? A little simplistic, don't you think? Especially in the light of the fact that it is patent attorneys who provide the written record, in the form of granted patents, that allows both private and public organizations to generate revenues from their discoveries. In an era of shrinking public funding of basic scientific research, these revenues are an indispensable tool to fund research.

Failing to protect such resources will cause unwary research organizations to incur significant losses. By denigrating the role of patent lawyers in protecting valuable intellectual property, you have done your readers a grave disservice.

Joseph T. Leone

*De Witt Ross & Stevens, 8000 Excelsior Drive,
Suite 401, Madison, Wisconsin 53717-1914, USA*

Who benefits from climate forecasts?

The effective and equitable dissemination of climate forecasts is as important and challenging as their accuracy. During El Niño 1997–98, Peruvian fisheries showed the need to understand forecast use and all parties' interests.

A. Pfaff, K. Broad and M. Glantz

El Niño often dominates world news about climate because of its widespread and frequently disastrous impacts. Our ability to predict El Niño is limited, but resources are committed to improving and communicating seasonal-to-interannual forecasts. Better forecasts will not necessarily help all, or even any, groups within an affected society, however. Attention must be focused on how to disseminate forecasts for maximum benefit.

A first issue is to study the use and non-use of forecasts in places affected by El Niño. This can help in designing dissemination strategies that address constraints on benefits such as limited access to forecasts and the distortion of forecast information. A second issue is to consider whose welfare counts as a 'benefit' among groups such as labour, industry and consumers; different regions; and the present and future. Although such choices are not made easier by analysis, they can be clarified by the delineation of different groups and their goals. These responsibilities may not belong to climate researchers themselves, but seem imperative for public agencies which fund such work and for those who promote the dissemination of forecasts.

Although El Niño was named by fishermen and described scientifically in Peru as early as 1892, it took the strong 1972–73 event and the collapse of the Peruvian anchovy fishery to draw attention to the potential impact of El Niño. And it was not until the 1982–83 event, which contributed to at least US\$13 billion of damage and 1,500 deaths, that El Niño's global nature was realized. Together with the growing belief that, given adequate data, forecasting was possible, these dramatic events stimulated the implementation of the Tropical Ocean–Global Atmosphere programme in 1985, and led to predictive models. Most researchers felt that such predictions were unreliable. But, after Mark Cane and Stephen Zebiak successfully forecast an El Niño in 1986, other groups began to release experimental forecasts to the public.

There were some successes. The Ethiopian government's advice to its farmers, after a forecast of drought in 1986–87, reduced the amount of food relief required. In northeast Brazil, a warning of drought in late 1991 prompted the state of Ceara to assist farmers to mitigate the effect on grain production, drinking water shortage and employment.

However, not all the news was good. In this Brazilian case, forecast credibility was questioned when the media spread early offi-



Caught out: El Niño can be disastrous for Peru's fishermen but many lack access to accurate forecasts.

cial forecasts but perhaps over-represented their certainty, and later official forecast updates were not spread. In Zimbabwe, a late 1997 forecast suggesting a more severe drought than occurred contributed to a reduction in planting and output. In Australia in 1997–98, a 'meteorological drought' (drop in rainfall) was correctly forecast, but the rains that fell came at a crucial time for crops, so that the 'agricultural drought' (drop in production) was less severe. Thus, those anticipating an El Niño as bad as the 1982–83 event lost out (for other examples, see M. H. Glantz, *Currents of Change: El Niño's Impact on Climate and Society*, Cambridge Univ. Press, 1996).

Today, El Niño forecasts are routinely disseminated, and climate institutions are expected to combine cutting-edge science with societal benefits. For example, the recently created International Research Institute for Climate Prediction, based at Columbia University in New York and the University of California at San Diego, has a mandate to improve forecasts and their applications, and is working with actual and potential 'end-users' of climate information.

Peruvian fisheries

Despite advance forecasts, Peru was severely affected by the extreme 1997–98 El Niño. Forecasts became widely available in March 1997 but, like many to follow, they underestimated the event's magnitude. Despite sizeable and relatively early disaster-prevention efforts, rain and floods affected thousands, crop yields fell and infrastructure was destroyed. The total fish catch fell steeply.

Commercial fishing had boomed from the mid-1950s until the early 1970s, when heavy fishing and the 1972–73 El Niño contributed to the collapse of the anchovy fishery. Only by the 1990s did catches recover. Today the sector represents more than 4% of the country's gross national product and in 1997 generated more than \$1 billion in foreign exchange earnings. It can be divided into artisanal and industrial groups, which fish different species using different methods, and are subject to different regulations.

The artisanal group consists of about 50,000 small-scale producers who fish or dive from small boats. Their catch is sold or consumed locally, or exported. Historically, they have been self-employed, with low socio-economic status and limited political influence. The industrial group employs more than 20,000 people, mainly in fishing fleets, fishmeal plants or canneries. Many more work in building and repairing ships, engines and nets. Fishermen's union power has disappeared as labour protection laws have weakened. Large firms have diversified into canned fish, agriculture, mining and other industries. They wield considerable influence at all levels of government. Because it has significant loans outstanding to fishing companies, the financial sector has considerable influence on the fishing sector.

During the 1997–98 El Niño, people responded both to the changing ocean environment and to a series of forecasts of the event's duration and magnitude. Industrial firms closed plants and canneries, moved fishing fleets south, and kept vessels idle. Illegal fishing increased, including industrial boats

pursuing fish into the zone reserved for artisanal fishing, which increased tension between the groups. Artisanal fishermen shifted to tropical species catchable only during El Niño conditions, but their revenues were limited by drops in market prices and climate-related interruptions of transport to markets.

Two other key groups, and potential forecast users, are the government and the financial sector. The Minister of Fisheries banned anchovy fishing relatively early, in April 1997. The ban was based on unusually high catches near shore and forecasts of a strong El Niño (but it was lifted after just a few days, reportedly due to industry pressure). For the rest of 1997 and well into 1998 catches plummeted. By September 1998, unemployment in Chimbote, the largest Peruvian fishing port, was so severe that food was distributed to thousands of fishermen and their families. The fishermen's welfare organization cut its employees' salaries and could not pay full benefits. In the financial sector, banks slashed lending based on climate forecasts, and re-financed many existing loans in anticipation of poor catches. Domestic effects on fishing families included increased criminality and reduced spending on basic education.

Outside the Net

The dominant medium for dissemination of probabilistic El Niño information in 1997–98 was the Internet, in English. The web is convenient, cheap and egalitarian, as it is widely available and easy to update. But although anyone in Peru with Internet access could view an El Niño website, in practice many (including artisanal fishermen and industrial employees) neither have access to a computer nor understand English. Also many Peruvians have little formal training regarding probabilistic information. Thus, climate forecasters unwittingly disadvantaged some groups in Peru compared with local competitors.

Second- or third-hand information can come through broadcast and print media. However, journalists often do not understand El Niño or probabilities, and 'simplify' the information. More effort is needed to target affected groups (perhaps via radio), and to make forecast information more useful, for example through public education or training of journalists.

Conflicting incentives

The Peruvian government has access to various forecasts of some skill, but scientists' limited ability to translate general predictions into coastal predictions for management of fish stocks leaves room for interpretation. As interpreters of forecasts, oceanographic and meteorological agencies often become advisers on decisions such as whether to declare a state of emergency. An emergency could freeze interest payments, helping companies desperate for debt relief. But in this event companies that wanted

loans or to sell bonds did not favour such a declaration. Given the latitude for interpretation, politics affects agencies' interpretations and whether governments listen to them. Such decisions can come down to relative influence. In 1997–98, different fishing companies tried to convince decision makers of the 'right' view to adopt of the intensity and duration of the El Niño event, and of its implications for fish stocks.

Institutions are made up of individuals. Individuals with access to climate information can demand payment to reveal it, blocking transmission. And it is rumoured that large fishing companies use tactics including job offers to distort official press releases. Even if incorrect, such beliefs undermine trust in official climate information.

Competition between institutions puts pressure on their employees to produce answers — now. In response, employees may remove caveats from press releases, injecting false certainty. With the media, they may rush to publicize sensationalized information. An article about a possible end to the El Niño event, in Peru's leading newspaper in late 1997, relied on the interpretation of a single satellite image, and conveyed unwarranted certainty. Media coverage of climate-related press releases also frequently draws its own conclusions on how mortality, crop yields or fish stocks will be affected.

Understanding the uses of forecasts could lead to better dissemination strategies. But possible conflicts may arise here too, since the pursuit of disseminators' interests (perhaps scientific progress or funding) may not maximize aid to those affected by El Niño. For example, some forecasters place all their latest data directly on the Internet (see, for instance, <http://nic.fb4.noaa.gov/products/predictions/experimental/bulletin/index.html>). However, the distribution of multiple, conflicting forecasts leads to confusion. Thus, for information to be of most use to those affected, it could be preferable to have passwords on websites for researchers, and a unified dissemination process, including quality control, for public climate information.

Better for everyone?

Stands are being taken on this controversial issue. For Peru, groups including the United Nations Food and Agriculture Organization and the World Bank have stressed the benefits of "responsible fishing", and advocated "sustainability". Other groups in Peru might disagree, as this could conflict with the fishing industry's focus on profit. From a sustainability viewpoint, one might define benefits as higher fish stocks. In that case, forecasts will not necessarily help. It may be obvious to supporters of sustainability that forecasts should trigger restrictions on fishing when stocks are under climate-related stress. However, forecasts were used by companies to anticipate fish migration patterns — pri-

vate use of El Niño forecasts contributed to increased fishing effort. Given a goal of sustainability, then, a better understanding of how locals use forecasts may suggest coupling their dissemination with tighter fishing regulations.

Others advocate equity within society, and the protection of workers during El Niño-related shocks to the fishing sector — obviously not a universal view. Again, the dissemination of forecasts will not necessarily yield benefits so defined, such as higher wages or employment. In the Peruvian case, when forecasts suggested low future profits, some fishing companies accelerated their seasonal downsizing.

The desire of forecasters and forecast providers for research progress and accessible climate information has contributed to an explosion of El Niño information. These goals are laudable and may even be widely accepted. However, unlimited dissemination may cause confusion. In the example we have discussed here, multiple, often conflicting forecasts in Peru have reduced the effectiveness of predictions. A research clearing-house might alleviate this problem.

Finally, meetings intended to produce consensus regional forecasts and communication between concerned citizens took place in Peru, sponsored in part by forecast providers. This is commendable, as is encouraging local groups to organize these forums. But particular local groups may have their own interests at heart. In Peru, some key agencies were excluded from full participation in a forum of this type.

General principles

The dissemination of seasonal-to-interannual climate forecasts can bring great benefits. However, to realize and maximize these, we must study the use of forecasts to help anticipate use patterns and increase the effectiveness of forecast dissemination. While politically challenging, it is also essential to make explicit whose 'benefits' count.

Although we have mainly discussed the Peruvian fishing sector during the 1997–98 El Niño, the lessons apply to the dissemination of El Niño forecasts to other regions and to economic sectors such as agriculture, disaster prevention and water management. They also apply to research choices made by climate scientists about the kinds of expertise and products to develop and provide. □

A. Pfaff is in the School of International and Public Affairs, Department of Economics, and Center for Environmental Research and Conservation, Columbia University, New York, New York 10027, USA. E-mail: ap196@columbia.edu. K. Broad is at the International Research Institute for Climate Prediction, Lamont-Doherty Earth Observatory of Columbia University, Palisades, New York 10964, USA. M. Glantz is at the Environmental and Social Impacts Group, National Center for Atmospheric Research, Box 3000, Boulder, Colorado 80307, USA.

Driving parents cuckoo

Douglas W. Mock

Do nestling birds signal their food needs honestly to their parents? A study of begging reed warblers shows that they do, but that parental sensitivities to vocal and visual cues can be exploited by an invading cuckoo chick.

The distinction between wants and needs lies at the heart of thinking on how dependent offspring communicate with their parents. Does a newborn send accurate and truthful information about its nutritional requirements, thereby helping its parents to allocate limited resources for the good of the whole brood? Or does it (unwittingly) practise deception, using 'psychological weaponry'¹ as a foil against its parents' physical superiority, and manipulating them

into skewing investment towards itself at the expense of its nest-mates?

A quarter of a century ago, R. L. Trivers' influential theory of parent-offspring conflict¹ challenged biologists to stop thinking of offspring merely as 'passive vessels'. Several theoretical approaches have since been used to model this parent-offspring conflict²⁻⁵ (Box 1). As is often the case, the choice of underlying assumptions can lead to several successful models — some in which honesty



Figure 1 Feed me now! — Cuckoo chick and its reed warbler 'parent'. Kilner *et al.*⁶ have shown that reed warblers use a combination of visual and vocal signals to work out how much food to bring to the brood. An invading cuckoo chick can fool the reed warblers into feeding it because, although its gape is not as large as that of a brood of warbler chicks, it makes as much noise as up to eight little warblers.

Box 1: Three fundamental concepts

The theory of parent-offspring conflict is most simply defined as a disparity between the optimal parental investment allocations for parents and offspring^{1-5,7,8,10}, with each offspring favouring skew towards itself, and parents more interested in equality for all. Each offspring is genetically only 50% identical-by-descent with even its full siblings, so it is, in effect, twice as 'related' to itself.

From that insight, Trivers¹ reasoned that a physically weak offspring might use 'psychological weaponry' — specifically, begging signals that inflated its true needs. This would con parents into the desired (and, for the parents, deleterious) skew. It is not clear what might constrain

such deceit, although various costs of the signals themselves have been proposed as forcing the whole system back to honest signalling^{2,3,9}.

Finally, the idea of the evolutionarily stable strategy^{13,14} is that many traits exist in a competitive milieu, with alternative versions of the same trait vying for supremacy. In the evolutionary context, the genes that underlie such traits ebb and flow until an equilibrium is reached that will not shift unless conditions affecting the replication rates of these genes change. At this equilibrium, the system is said to be evolutionarily stable, and the unbeatable winner is the evolutionarily stable strategy. D. W. M.

of the offspring is the only evolutionarily stable strategy^{2,3}, others where various 'compromise' outcomes are possible^{4,5}.

Meanwhile, field biologists have been busy as well and, on page 667 of this issue, Kilner and colleagues⁶ describe a two-pronged research programme into honesty and deception among avian nestlings. First, they worked out how the begging signals of nestling reed warblers (*Acrocephalus scirpaceus*) convey information that seems to be basically honest (at least, as far as the current surrogates of 'need' can detect) to the parents. The authors found that warbler chicks increase their gapes and calling rates both as they grow older and as the interval increases since they were last fed. Similarly, chicks that naturally show large gapes and high calling rates consume more food than their less enthusiastic siblings when allowed all the food they want. I leave the reader to make the crucial — and extremely difficult — step of accepting gluttony as a legitimate measure of physiological (much less evolutionary) need. But there is no indication that reed warbler nestlings rip off their parents^{7,8}. Perhaps they are limited by the costliness of their displays, as predicted by honest-signal theory^{2,3,9}. Alternatively, maybe the gap between the evolutionary interests of the parents and offspring is trivial, or even non-existent^{5,7,8,10}.

The second research prong, though, blows away all nuance. When the parent is a reed warbler and the offspring is a common cuckoo chick (*Cuculus canorus*) whose mother inserted its egg (thoughtfully pre-warmed in her oviduct for an extra day to accelerate hatching) into the warbler nest, all congruencies of Darwinian fitness are out of the window — this is war! The cuckoo hatching evicts its nest-mates, hatched or otherwise, then accepts all the food that the two warbler parents deliver while growing into a grotesque behemoth¹¹ (Fig. 1). Kilner *et al.*⁶ now show that cuckoos produce a begging signal that evades the parent warblers' radar. This signal fits overwhelmingly with what the warblers want to hear, if only imperfectly with what they want to see.

Nestling reed warblers offer two relevant signal components, one visual (the gape area of the entire brood) and one vocal (a chorus of 'si...si...si...si' calls). These operate more or less independently, and account for different components of the total variance in parental deliveries. The tandem of signals conveys more accurate information about offspring condition than either could alone, so the display is legitimately considered a 'multiple' signal⁹. Parent warblers must have some mental trick for integrating these signal components and adjusting how much effort they put into feeding their brood. For human readers, the 'simple behavioural rule' of parent warblers is rendered as a regression equation that accounts for an impressive 62.7% of the variance in parental response.

Faced with this plethora of warbler parent predilections, the cuckoo chick has a serious problem that it must solve. Despite its gargantuan size it has a relatively small beak, so it cannot match the visual stimulus of four gaping warbler nestlings. But Kilner *et al.*⁶ have found that warbler parents are sensitive to an array of visual and vocal signal combinations. This leaves room for a bit of bait-and-switch, wherein understatement in one sensory signal can be made up by overstatement in another. So, the cuckoo's beak deficiency is compensated for by vocal wizardry that not only matches the total calling rate of four warbler nestlings (at the age of one week), but rises so fast as the cuckoo grows that it soon sounds like eight little warblers! Interestingly, the total stimulus package — the voice of eight chicks combined with the gape of roughly two — inspires parents to their normal work ethic, and they bring enough food for four warbler chicks to the single interloper. By establishing that warbler parents use the same integration rule when dealing with true offspring and cuckoos alike, Kilner *et al.* solved a long-standing ethological puzzle over why nestling cuckoos have such extraordinary calls. Of course, the deficiency in the visual signal was caused by the cuckoo's habit of throwing out the host young in the first place, presumably opting for sole access to the delivered food in lieu of retaining some help with the signalling problem.

But why does the cuckoo stop there? Why not hustle the expendable warbler parents into delivering food for five, six or seven? With no shared genetic interests, why not drive them into an early grave? Kilner *et al.* suggest two good reasons. First, the cuckoo needs its 'parents' for ten days more than the warbler offspring would have needed them, so there may be an advantage in allowing the parents to pace themselves and work at a normal level. According to this view, greed and pragmatism are in balance. Alternatively, the cuckoo may have hit a ceiling and, if it cannot sing more rapidly or gape more widely, perhaps it cannot exploit the warbler parents further. For that matter, perhaps the cuckoo chick already gets as much food as it can process — a point that might easily be checked by seeing whether cuckoos of the same species raised by different hosts grow faster in the nests of rapidly delivering hosts and more slowly in meaner circumstances.

So it is that, in the evolutionary arms race between warbler and cuckoo, the parasite has decoded the host's integration rule, adapting what it can offer to cover what it cannot. This phenomenon of pandering to pre-existing sensitivities is known in other contexts, ranging from sexual signals in mites, fishes and frogs¹² to faux sexual signals in *Homo sapiens* (toupees, breast implants, elevator shoes and codpieces). In the cuckoo–host system, it also raises a wonderful ques-

tion about specificity. Common cuckoos parasitize many species of songbird — some of them apparently for millennia, such that the cuckoo has evolved extraordinary egg mimicry. So, can we expect to see similar communication adaptations in sub-populations of cuckoos that parasitize particular hosts for the multiple-signal integration rules of each host? □

Douglas W. Mock is in the Department of Zoology, University of Oklahoma, 730 Van Vleet Oval, Norman, Oklahoma 73019, USA.
e-mail: dmock@ou.edu

1. Trivers, R. L. *Am. Zool.* **14**, 249–264 (1974).
2. Godfray, H. C. J. *Nature* **352**, 328–330 (1991).
3. Godfray, H. C. J. *Nature* **376**, 133–138 (1995).

Condensed-matter physics

A nanotube laboratory

David H. Cobden

Fifty years ago and more, physicists fused beams of particles guided through vacuum chambers to investigate fundamental issues such as wave–particle duality and the scattering of particles from one another. More recently, they have tended to prefer to do similar experiments in cleverly constructed condensed-matter systems. This is because electrons and photons often pass through solids or liquids very much like they do through a vacuum, but with subtle, varied and highly controllable modifications to their behaviour. The latest trend is to use not bulk materials but individual, appropriately formed molecules to form the experimental apparatus. The potential for doing basic physics experiments using a single molecule is nicely illustrated by Bachtold *et al.* on page 673 of this issue¹, where they report observation of the Aharonov–Bohm effect in a single carbon nanotube.

In the classical world of Newton's laws and Maxwell's equations, an electron is deflected by the electromagnetic fields it moves through, but is completely indifferent to the fields everywhere else in the Universe.

4. Parker, G. A. & Macnair, M. *Anim. Behav.* **27**, 1210–1235 (1979).
5. Mock, D. W. & Parker, G. A. *The Evolution of Sibling Rivalry* (Oxford Univ. Press, 1997).
6. Kilner, R. M., Noble, D. G. & Davies, N. B. *Nature* **397**, 667–672 (1999).
7. Godfray, H. C. J. *Am. Nat.* **146**, 1–24 (1995).
8. Mock, D. W. & Forbes, L. S. *Trends Ecol. Evol.* **7**, 409–413 (1992).
9. Johnstone, R. A. in *Behavioural Ecology: An Evolutionary Approach* (eds Krebs, J. R. & Davies, N. B.) 155–178 (Blackwell, Oxford, 1997).
10. Mock, D. W. & Parker, G. A. *Anim. Behav.* **56**, 1–10 (1998).
11. Davies, N. B., Kilner, R. M. & Noble, D. G. *Proc. R. Soc. Lond. B* **265**, 673–678 (1998).
12. Ryan, M. J. & Rand, A. S. *Phil. Trans. R. Soc. Lond. B* **340**, 187–195 (1993).
13. Maynard Smith, J. *Evolution and the Theory of Games* (Cambridge Univ. Press, 1982).
14. Dugatkin, L. D. & Reeve, H. K. (eds) *Game Theory and Animal Behavior* (Oxford Univ. Press, New York, 1998).

In 1959, Aharonov and Bohm² pointed out that things are different in the world of quantum mechanics. They proposed an experiment in which a beam of electrons is split into two and then recombined in the vicinity of a magnetic field (Fig. 1a). Thanks to the wave-like side of an electron's nature, an interference pattern is formed on a screen, just as in the Young's slits experiment with light. Curiously, the magnetic field enters Schrödinger's wave equation only rather indirectly, through the so-called vector potential, which is an integral of the magnetic field over space. The result is that the positions of the interference fringes on the screen are sensitive not to the magnetic field along paths 1 and 2, but rather to the total magnetic flux through the region between the two paths. The positions of the fringes cycle each time the flux is changed by the quantum unit h/e . In fact, the magnetic field can be completely confined to a solenoid, so that the electrons never come near the magnetic field which influences them, emphasizing the difference from the classical situation.

It was not long before the Aharonov–Bohm effect was observed³, using the

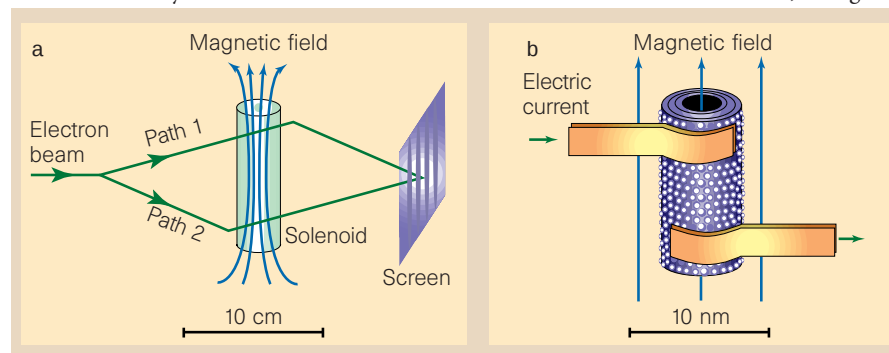


Figure 1 Observing the Aharonov–Bohm effect — the quantum mechanical influence of a magnetic field on the motion of electrons — at very different scales. **a**, The setup in the traditional physics laboratory. **b**, The setup in a molecular laboratory, such as the system described by Bachtold *et al.*¹. Here the experiment is performed using a single molecule, a multiwalled carbon nanotube.

well-established techniques for manipulating electron beams in a vacuum. Two decades later the new field of mesoscopic physics arose, in which pieces of metal or semiconductor were made so small and cold that the conduction electrons moved around in them as coherent waves. Fundamental to mesoscopic systems is the interference between a host of different paths available to electrons passing from one point to another. In a magnetic field the Aharonov–Bohm effect should occur between every pair of possible paths, with important but generally rather messy consequences, because the paths are often very complicated and numerous.

It was predicted⁴ however that, for the particular geometry of a small, thin-walled hollow metallic cylinder in an external magnetic field (Fig. 1b), the Aharonov–Bohm effect should cause the electrical conductance to oscillate with the magnetic flux through the cylinder's bore. When researchers evaporated a normal metal film onto a two-micrometre-thick insulating fibre they found exactly this behaviour⁵. Moreover, since then the Aharonov–Bohm effect has become thoroughly established as a basic principle in the physics of mesoscopic systems (many of which, incidentally, are created by drawing patterns with good old electron beams).

Nearly two decades further on, mesoscopic devices have now become 'nanostructures', whose size scales reach down to the molecular level, and which sometimes even employ individual molecules at their heart. Rapidly becoming the archetype of these is a particular class of wire-like molecules with extraordinary electrical properties — the carbon nanotubes⁶. The simplest nanotubes are long, flexible, hollow cylinders, each like a drinking straw rolled from a single graphite layer, where the atoms are arranged in a hexagonal lattice. Because these tubes are so narrow (around 1.5 nm in diameter), electrons are quantum mechanically restricted to move only parallel to the tube axis. In the past couple of years this has opened up a real-life laboratory for one-dimensional physics^{7,8}, long the preserve of theorists. Another, more common type of nanotube contains many coaxial graphitic cylinders. In these thicker (typically 10-nanometre or more) multiwalled tubes, the electrons can move relatively freely over the outer cylindrical surface — just the thing for observing the Aharonov–Bohm effect.

Bachtold *et al.*¹ attached metal leads to individual multiwalled tubes, and found conductance oscillations consistent with the metallic cylinder theory. Their results dramatically demonstrate the potential for science in such nanostructures, where physics and chemistry merge. The electron-beam paths in the Aharonov–Bohm experiment have effectively been replaced by molecular

orbitals. Bachtold *et al.* were also surprised to find extra oscillations with a smaller period, corresponding to electron paths that revolve several times around the cylinder before interfering. Although there is no specific explanation for these oscillations yet, they could well represent the underlying nature of the orbitals in the nanotube, such as might be caused by a built-in twist in the graphite lattice. This illustrates the likelihood that, thanks to the infinitely rich nature of molecular orbitals, future molecular-scale experiments will lead to new science that goes far beyond the domain of the old-fashioned macroscopic laboratories.

It was not by coincidence that this seven-orders-of-magnitude reduction in the size of physics experiments came over the same period that electric valves, arc lamps and cathode ray tubes were replaced by transistors, solid-state lasers and flat-panel dis-

plays. The technology of nanostructures is driven largely by the urge to make ever smaller and faster electronics⁹. However, scientists are both pushing and riding the wave, and physicists will be busy for a long while exploring and exploiting their new nanometre-sized laboratories. □

David H. Cobden is at The Niels Bohr Institute, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen, Denmark.
e-mail: cobden@fys.ku.dk

1. Bachtold, A. *et al.* *Nature* **397**, 673–675 (1999).
2. Aharonov, Y. & Bohm, D. *Phys. Rev.* **115**, 485–491 (1959).
3. Chambers R. G. *Phys. Rev. Lett.* **5**, 3–5 (1960).
4. Al'tshuler, B. L., Aronov, A. G. & Spivak, B. Z. *JETP Lett.* **33**, 94–97 (1981).
5. Sharvin, D. Y. & Sharvin, Y. V. *JETP Lett.* **34**, 273–275 (1982).
6. McEuen, P. L. *Nature News & Views* **393**, 15–17 (1998).
7. Tans, S. J., Devoret, M. H., Groeneveld, R. J. A. & Dekker, C. *Nature* **394**, 761–764 (1998).
8. Bockrath, M. *et al.* *Nature* **397**, 598–601 (1999).
9. Sohn, L. L. *Nature News & Views* **394**, 131–132 (1998).

Development

Hox proteins reach out round DNA

Matthew P. Scott

It might be misquoted of Hox genes that “in the field of development, never was so much owed by so many tissues to so few genes”. All animals have Hox genes, and the multifarious roles of Hox proteins in building vital organs such as brain, muscle, bone and gut endow them with honorary status among the transcription factors, both in controlling the fates of different cells and in arranging them into working structures. Valuable insights into how Hox proteins work are offered by two new papers, one by Passner *et al.*¹ on page 714 of this issue and the other by Piper *et al.*² in last week's *Cell*, which each describe the structure of a Hox protein on its DNA-binding site in combination with a cofactor protein.

Over a century ago, Bateson³ described homeotic transformations as a change in one body part into the likeness of another, broad-mindedly including misplaced insect appendages and transformations of mammalian vertebrae. The vertebrae pattern switches were both anterior to posterior and vice versa, which are the sorts of changes that are now known to be caused by ectopic or by reduced Hox gene function.

Hox genes encode transcription factors that coordinate the expression of genes necessary for the implementation of different developmental pathways, so they can generate some bizarre phenotypes when mutated: for example, *Ultrabithorax* mutations in the fruitfly *Drosophila* create an extra set of wings, and *Antennapedia* mutants grow legs in place of antennae; in the nematode *Caenorhabditis elegans*, Hox mutations reroute migrating cells; in mouse, they convert one type of vertebra into another or change

the pattern of hindbrain development; and in humans, they cause fusions and duplications in hand and foot digits.

Hox genes are a specific subset of the hundreds of homeobox genes⁴. The homeobox is a 180-base-pair DNA sequence that encodes a 60-amino-acid DNA-binding domain, the homeodomain. Hox genes are distinguished from other homeobox genes both by their more closely related homeodomain sequences and by their clustered positions on chromosomes. Four clusters of mammalian Hox genes, 39 genes in all, are thought to be derived from one ancestral cluster — in flies, the single ancestral cluster evidently split into two parts. Within each fly or mammalian cluster, Hox genes are conveniently organized according to the body region they affect: at one end of the cluster, genes are found that are involved in anterior (head) development; at the other end, the genes work for more posterior structures. In between, genes are arranged roughly in the order of their influence along the body, proceeding from head to tail. Each mammalian complex is designated as A–D, and each gene within the complex by a paralogue number, 1–13: paralogue 1 genes act on the most anterior regions of the body axis, and paralogue 13 genes on the most posterior ones. Genes that have the same paralogue number but are located in different mammalian clusters are highly related and sometimes partially redundant in function.

The *Ultrabithorax* (*Ubx*) protein studied by Passner *et al.*¹ is produced and functional in the posterior thoracic and anterior abdominal segments of *Drosophila*⁵. *Ubx* mutations lead to posterior-to-central tho-

rax transformations, giving rise to Lewis's famous four-winged fly⁵, and to abdomen-to-thorax transformations. The human HoxB-1 protein, the subject of the study by Piper *et al.*², is required for normal development of rhombomere 4 of the hindbrain⁶, where *hoxb-1* is expressed. *hoxb-1* is most related to the *Drosophila* 'head' gene *labial*, and is therefore in a different class from *Ubx*, which is most closely related to Hox6–8.

Hox protein products of both *Ubx* and *hoxb-1* interact functionally with another group of homeodomain proteins called Extradenticle (Exd) in flies and Pbx in mammals. These proteins enhance the specificity and affinity of DNA binding by Hox proteins⁷. The *pbx1* gene was first found at chromosome breakpoints associated with human leukaemia. Mutations in *exd* cause diverse homeotic transformations, and yet *exd* does not transcriptionally regulate most Hox genes. Instead, Exd cooperates with Hox proteins to bind to DNA, so loss of Exd function leads to a compound phenotype incorporating multiple Hox phenotypes. Similarly, Pbx1 has been found to cooperate with Hox proteins in, for example, *hoxb-1* autoregulation⁸. Thus, Exd/Pbx1 proteins act as cofactors that increase the DNA-sequence specificity of Hox proteins.

The two new crystal structures that have been solved for the Ubx–Exd–DNA and HoxB-1–Pbx1–DNA complexes^{1,2} both display the three α -helices that are characteristic of homeodomain structures. The DNA contacts for Exd and Pbx1 differ from those of most homeodomains, however, and may be weaker or less discriminating. The cooperative binding between the Hox and Exd/Pbx1 proteins is due to a hexapeptide outside and amino-terminal to the homeodomain. The crystal structures reveal how the individual components cooperate in binding to one another: the two homeodomains bind to opposite sides of the double helix of DNA. The Hox protein hexapeptide, which is on the end of a long linker arm, reaches around to Exd or Pbx1 on the other side, inserting itself into a special pocket in Exd or Pbx1 which is created partly by a tripeptide loop peculiar to the Exd/Pbx1 group of homeodomains. Because both homeodomains make sequence-specific contacts with bases (sometimes the same bases) and with backbone atoms, the specificity and binding affinity of the whole complex are enhanced.

Remarkably, a region of Pbx1 that is outside the homeodomain on the carboxy-terminal side, and which is essential for the cooperative binding of Pbx1 and HoxB-1, does not touch either HoxB-1 or the DNA. Instead, it forms a fourth helix that packs against the third helix, which is the most important determinant of sequence specificity. This fourth helix is presumed to stabilize or modify the helix-3 configuration in

order to impart greater DNA-binding strength and to hold helix-3 in an optimal contact position for insertion of the hexapeptide. In addition, the DNA is bent by about 10° by both Pbx1 and HoxB-1, which may influence the strength of binding.

What are the implications of the features revealed in these structures? First, Hox proteins of different paralogue groups associate in quite a similar way with Exd/Pbx proteins. Second, the basis for sequence specificity is now clearer than when Hox proteins alone were tested on DNA, but additional genuine target-gene *cis*-regulatory sequences need to be identified before we can find out how the combinatorial sequence recognition is actually employed. Third, the different actions of various paralogue groups could be accomplished by association with proteins other than Exd/Pbx. Some paralogue groups have a hexapeptide sequence of Y/F-P-W-M-K/R (single-letter amino-acid code), whereas others contain a tryptophan (W) residue in a different context⁴. Hexapeptides of different paralogue groups have distinct sequences, paralogue groups have characteristic linker lengths between hexapeptide and homeodomain, and some paralogue groups have characteristic residues carboxy-terminal to the homeodomain⁹. Each type of hexapeptide could insert into a different cofactor if more proteins like Exd/Pbx exist. One candidate is the Meis/Hth protein, a cofactor for Exd/Pbx^{10–12}. Alternatively, the linker length between the hexapeptide and the homeodomain could affect which target-gene sequence can be bound by the combination of a Hox protein and Exd or Pbx. A stretched

or distorted linker could alter the insertion orientation of the hexapeptide. Fourth, the sequence conservation found among some Hox paralogues carboxy-terminal to the homeodomain may indicate that such regions have an indirect role in DNA binding, as in the case of Pbx1.

Satisfaction comes from seeing universals emerge, such as structural similarities between homeodomains and bacterial helix–turn–helix proteins and yeast mating-type homeodomain proteins — but at the same time, the new Hox–Exd/Pbx structures are rich in novelty. Bateson would be pleased. □

Matthew P. Scott is in the Departments of Developmental Biology and Genetics, Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, California 94305, USA.

e-mail: scott@cmgm.stanford.edu

1. Passner, J. M., Ryoo, H. D., Shen, L., Mann, R. S. & Aggarwal, A. K. *Nature* **397**, 714–719 (1999).
2. Piper, D. E., Batchelor, A. H., Chang, C.-P., Cleary, M. L. & Wolberger, C. *Cell* **96**, 587–597 (1999).
3. Bateson, W. *Materials for the Study of Variation* (Macmillan, London, 1894).
4. Duboule, D. *Guidebook to the Homeobox Genes* (Oxford Univ. Press, 1994).
5. Lewis, E. B. *Nature* **276**, 565–570 (1978).
6. Studer, M., Lumsden, A., Ariza-McNaughton, L., Bradley, A. & Krumlauf, R. *Nature* **384**, 630–634 (1996).
7. Mann, R. S. & Chan, S. K. *Trends Genet.* **12**, 258–262 (1996).
8. Popper, H. *et al.* *Cell* **81**, 1031–1042 (1995).
9. Sharkey, M., Graba, Y. & Scott, M. P. *Trends Genet.* **13**, 145–151 (1997).
10. Chang, C. P. *et al.* *Mol. Cell. Biol.* **17**, 5679–5687 (1997).
11. Rieckhof, G. E., Casares, F., Ryoo, H. D., Abu-Shaar, M. & Mann, R. S. *Cell* **91**, 171–183 (1997).
12. Berthelsen, J., Zappavigna, V., Ferretti, E., Mavilio, F. & Blasi, F. *EMBO J.* **17**, 1434–1445 (1998).

Bioenergetics

One price to run, swim or fly?

R. McNeill Alexander

It has long been accepted that it is metabolically cheaper for animals to fly than to run, and cheaper still for them to swim¹. Measurements of oxygen consumption have shown, in comparisons between animals of equal mass, that the energy used by a mammal to run one kilometre is enough to enable a bird to fly about two kilometres, or a fish to swim up to ten kilometres. But in a recent paper in *Philosophical Transactions of the Royal Society*, Williams² repeats the analysis for mammals alone and reaches the surprising conclusion that their three types of locomotion are almost equally priced. Thus, a 100-kg seal needs as much energy to swim a kilometre as a 100-kg pony would need to run a kilometre, and a 1-kg fruitbat needs only a little less energy to fly a kilometre than a mongoose of the same mass needs to run it².

The energy cost of animal locomotion is expressed as the cost of transport (energy used)/(body mass $\times$ distance travelled). It

may be defined either to include all the metabolic energy used on the journey (total cost of transport) or only the extra energy, excluding that which would have been used if the animal had been resting (net cost).

Williams has compared total costs of transport for marine mammals with different degrees of aquatic specialization, ranging from muskrats and sea otters, which have bodies and paws like those of their fully terrestrial relatives, to sea lions with streamlined bodies and flippers, and even whales, which have evolved tail flukes for efficient swimming. She finds that swimming is very expensive for semi-aquatic mammals: a 20-kg sea otter swimming at the surface uses five times as much energy per kilometre as a sea lion of the same mass swimming under water. For this reason, semi-aquatic mammals are limited to low speeds. Sea otters are slower than sea lions, and dolphins are five times as fast, over 50 metres, as world-record-holding human

NATURE

100 YEARS AGO

The last letters received from Mr. John S. Budgett, who has been sent out on a scientific mission to the Gambia by the Zoological Society of London, are dated from the Nianimaru ... about thirty miles below McCarthy Island, January 23. They announce that Mr. Budgett, who was in excellent health, was busily engaged in collecting fishes and birds. Of *Polypterus*, one of the objects of his special inquiries, he had obtained some large specimens, one of which was found to contain large ova. The Manatee (*Manatus senegalensis*) had been ascertained to ascend the river thus far.

From a report in the *Adelaide Chronicle*, it seems that the iguana lizard, hitherto supposed to be perfectly harmless, except in poultry yards, has been found to be the slaughterer of lambs. Several sheep-owners have caught the iguana in the act of killing lambs, so there can be no doubt that this reptile must now be classed amongst the enemies of the pastoralists. The scarcity of opossums has probably driven the iguana to attack lambs for food. Pastoralists ... say that even very small iguanas will attack a lamb, and they further state that any lamb so bitten will not recover.
From *Nature* 23 February 1899.

50 YEARS AGO

It is commonly stated ... that the energy of muscular contraction is derived in the first instance from the breaking of the terminal energy-rich phosphate bond of adenosine triphosphate. Why not try to find out whether it really is, not in muscle extracts which cannot contract but in muscles which can? ... Prof. T. B. L. Webster recently provided an epigram to be posted in my laboratory: 'Υποθεσις ελεγχουζ μη αποδεχομενη ουχ εξουτεα', which means that hypotheses should not be put forward that do not admit of test. The hypothesis that adenosine triphosphate breaks down during the contraction of a living muscle ought not to be left in that category, but only biochemists can remove it. – A. V. Hill
From *Nature* 26 February 1949.

Many more extracts like these can be found in *A Bedside Nature: Genius and Eccentricity in Science, 1869–1953*, a 266-page book edited by Walter Gratzer. Contact Lisa O'Rourke. e-mail: l.orourke@nature.com

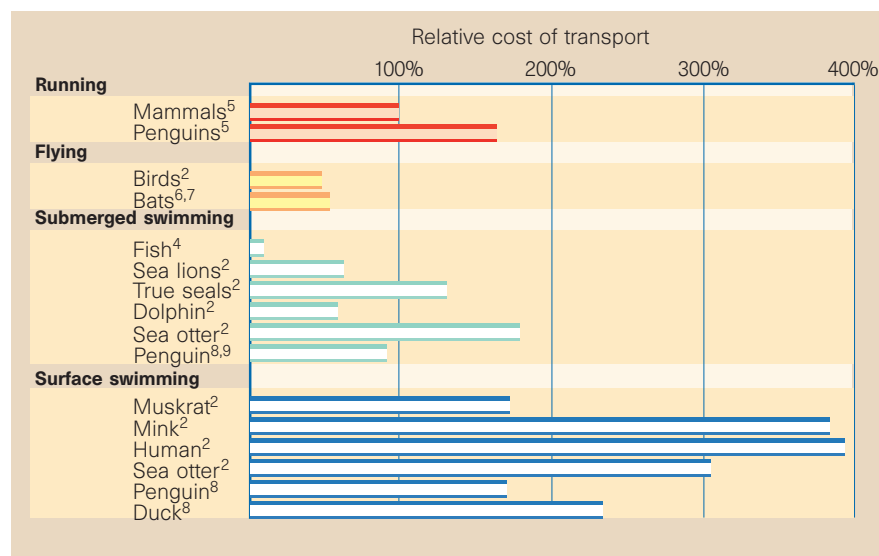


Figure 1 Total costs of transport for animals running, swimming and flying, as percentages of the cost for a running mammal of the same mass. The values used for running are the asymptotic values at high speeds (the 'slopes' of ref. 5).

swimmers³. Indeed, dolphins can swim as fast as human sprinters can run.

Figure 1 compares some total costs of transport with costs for running mammals. This confirms that endothermic submerged swimmers, penguins as well as marine mammals, have costs of transport similar to those of running mammals, and much higher than those of fish. It also shows that surface swimmers face much higher costs. However, with some additional data not used by Williams², it seems to show that costs for flying bats are only a little higher than for birds.

The similarity of costs of transport for running and swimming mammals must surely be coincidental, because work is used, in these two modes of locomotion, in quite different ways. If so, why is swimming much more expensive for mammals and penguins than it is for fish? The answer seems to be that mammals and birds have much higher resting metabolic rates than fishes of similar size, so have to swim much faster to minimize total cost of transport. (The mechanical power required to propel a body through water at speed v is about proportional to v^3 (ref. 4), so the metabolic rate can be expected to be $(R + kv^3)$, where R is the resting rate and k is a constant. The energy required per

unit distance is $(R + kv^3)/v$, which has its minimum value when $v = (R/2k)^{0.33}$. Thus, a large resting metabolic rate implies a high optimum speed and a high cost of transport.) The speeds required to minimize total cost of transport are about 1.8 times as high for seals, dolphins and penguins as for fish⁴.

For sea otters and penguins, we know the costs of transport for swimming both at the surface and submerged (Fig. 1). In both cases, surface swimming is much the more costly, which we would expect as a surface swimmer has to push a bow wave along. Tests on inert bodies towed along the surface of water have recorded up to five times as much drag as when they were towed at the same speed, deeply submerged⁴. However, comparison of a sea otter with a penguin, either both at the surface or both submerged, shows that the penguin performs much the better. Also, a penguin at the surface has a much lower cost of transport than a duck of equal mass at the surface. The penguin benefits from its more streamlined body, and from hydrofoil flippers that can push on larger masses of water than webbed feet are able to, making higher propeller efficiencies



Figure 2 Laid back: a Californian sea otter keeping down the energy cost of locomotion.

possible⁴. Sea lions and dolphins have similar advantages. The penalty for penguins is that their adaptations for swimming make them ungainly waddlers on land, with much higher costs of transport than for typical mammals of equal mass⁴. Economical running and economical swimming seem to be incompatible. □

R. McNeill Alexander is in the School of Biology, University of Leeds, Leeds LS2 9JT, UK.

Neurobiology

Sorting out the neuron

Peter J. Hollenbeck and Gordon Ruthel

From bacteria to algae to eggs, all cells have some degree of asymmetry — or polarization — in their structure and function. Functional polarization relies on specific macromolecules being sorted to particular sites in the cell, and the challenge is to work out how cells sort and locally retain macromolecules. Perhaps nowhere is this so difficult as in the case of proteins in the cell membrane, because their ability to diffuse in the plane of the membrane should defy the cell's attempts to position them. Some cells (such as epithelia) have obvious physical barriers to free diffusion of membrane components¹. But although nature's most polarized cells, neurons, lack such barriers, certain proteins are clearly restricted to specific regions of the plasma membrane. How can this be?

On page 698 of this issue, Winckler and co-workers² shed welcome new light on this conundrum. They have pinpointed a barrier to protein diffusion in the membrane of cultured neurons, and characterized it using direct and quantitative physical methods. Their results indicate that membrane proteins interact with a local, specialized region of the protein cytoskeleton underneath the membrane. This region impedes diffusion of the membrane proteins, allowing the neuron to maintain separate protein domains within its plasma membrane.

The neuron can be divided into three functional domains (Fig. 1). First is the soma, where the nucleus, endoplasmic reticulum and Golgi apparatus reside, and from which the other domains extend. Second comes the axon, a highly elongated, specialized process that conveys information to the periphery. Finally, there are the dendrites, shorter processes that receive and integrate information. Although there seems to be no significant barrier to lipid diffusion among the regions of plasma membrane that overlie these domains³, many membrane proteins are segregated either to the soma and dendrites (the 'somatodendritic domain') or to the axon. So, it was assumed that any barrier to protein diffusion must lie between the axon and the soma. Winckler *et al.*² have now identified a boundary between axonal and

- Schmidt-Nielsen, K. *Science* **177**, 222–228 (1972).
- Williams, T. M. *Phil. Trans. R. Soc. B* **354**, 193–201 (1999).
- Alexander, R. McN. *The Human Machine* (Natural History Museum Publications, London, 1992).
- Videler, J. J. *Fish Swimming* (Chapman & Hall, London, 1993).
- Taylor, C. R., Heglund, N. C. & Maloij, G. M. O. *J. Exp. Biol.* **97**, 1–21 (1982).
- Carpenter, R. E. J. *Exp. Biol.* **114**, 619–647 (1985).
- Thomas, S. P. J. *Exp. Biol.* **63**, 273–293 (1975).
- Baudinette, R. V. & Gill, P. J. *Comp. Physiol. B* **155**, 373–380 (1985).
- Culik, B. M., Wilson, R. P. & Bannasch, R. J. *Exp. Biol.* **197**, 65–78 (1994).

somatodendritic membrane proteins that lies 20–90 μm from the soma in a region of the axon called the initial segment.

To find out how this boundary works, Winckler and colleagues applied tiny silica beads to the axonal surface. The beads were coated either with antibodies that bind the extracellular portion of specific membrane proteins, or with lectins, which bind most membrane proteins. The authors then used an optical trap, or 'laser tweezers', to manipulate the beads — they dragged the membrane proteins laterally along the axon, with-

in the plane of the membrane, until they detached from the trap. In all cases, the beads and their attached membrane proteins could be moved over ten times further in the distal axon than they could in the initial segment, indicating that the initial segment impedes lateral mobility of proteins in the membrane. This was true both for L1, a protein that spans the plasma membrane, and for Thy1, which is linked only to lipids in the outer leaflet of the membrane. Furthermore, the reduced tractability of lectin-coated beads indicated that many membrane proteins experience the same limits to diffusion in the initial segment.

Winckler *et al.* next probed the mechanism by which the initial segment limits diffusion. They disrupted the cytoskeleton that underlies the plasma membrane and found that this treatment eliminated the diffusion barrier of the initial segment — proteins normally confined to the somatodendritic or axonal membrane gradually diffused out of their proper domains, and membrane proteins coupled to beads could be dragged freely through the initial segment using laser tweezers. Based on these data, and by analogy with other cell types⁴, an attractive mechanism for restricting diffusion would be a tethering interaction between membrane

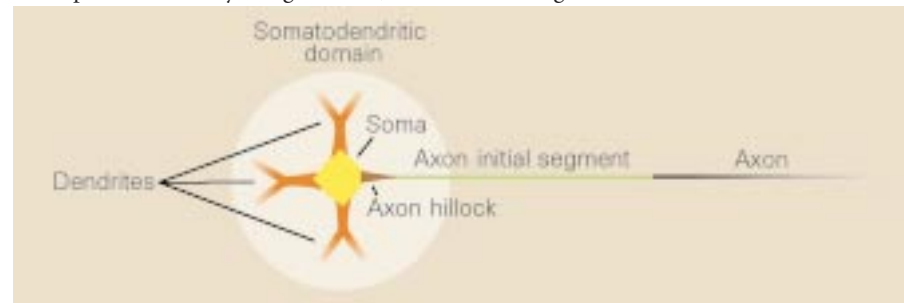


Figure 1 A typical neuron and its components. The cell body (soma) gives rise to several dendrites to form the somatodendritic domain, which has a different array of surface proteins to the axonal domain. The axon arises from the soma at the axon hillock, and the first 60–90 μm of the axon is referred to as the initial segment.

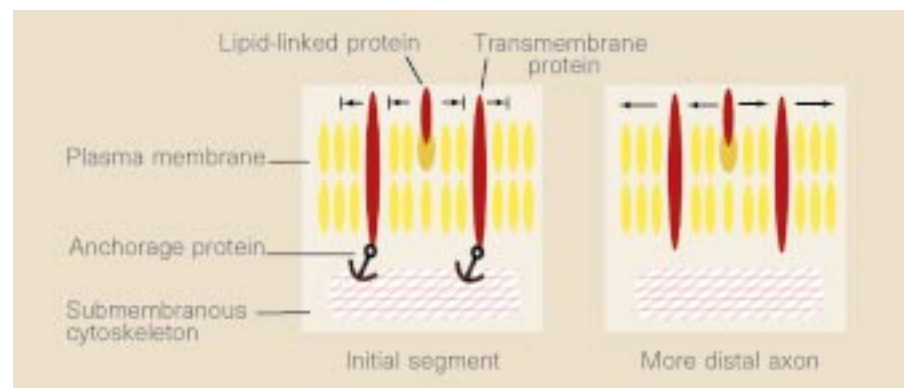


Figure 2 Movement of proteins in various regions of the neuron. Proteins may span the plasma membrane or be linked to lipids within the outer leaflet. Winckler *et al.*² have shown that in distal regions of the membrane (right), proteins move freely within the membrane. But in the initial segment (left), movement is hindered. The authors propose that transmembrane proteins are attached to anchoring proteins in the specialized protein skeleton beneath the membrane in this region. The stationary transmembrane proteins may, in turn, constrain free movement of other proteins in the membrane.

proteins and the cytoskeleton (Fig. 2). Consistent with this, when the authors used a detergent to extract free membrane constituents, transmembrane proteins were removed from the whole axon except the initial segment, where they were retained.

The entire neuron has a submembranous cytoskeleton, so what differentiates the initial segment as a site for interaction with membrane proteins? Winckler *et al.* found that the initial segment is enriched for two submembranous proteins. One of these — ankyrin G (relative molecular mass, 480/270K) — directly interacts with the cytoplasmic domains of many neuronal transmembrane proteins⁵. Ankyrin G has also been implicated in restricting specific transmembrane proteins to nodes of Ranvier⁶ (gaps in the lipid myelin sheath that surrounds the axon). Maybe it has a similar function in the initial segment, contributing to a functional boundary by binding to proteins that would otherwise diffuse past.

Although interaction with the cytoskeleton could hinder diffusion of transmembrane proteins like L1, a different mechanism must be proposed for proteins like Thy1. These proteins are segregated by the initial segment, but they do not span the membrane and can be easily extracted from it. Perhaps, as the authors suggest, the sheer density of transmembrane proteins restrained in the initial segment creates a traffic jam that indirectly impedes diffusion of peripheral proteins such as Thy1. This type of mechanism would be supported by

the tendency of such lipid-linked proteins to associate with larger, macromolecular complexes in the membrane⁷.

For the initial segment to be an effective barrier, membrane proteins must first be directed to their appropriate compartments⁸. Working out how these proteins are initially sorted will be at least as important as understanding how molecular segregation is maintained after sorting. Neurons lack a diffusion barrier to separate membrane components of the soma and dendrites, even though many cytoplasmic constituents are segregated between these compartments. This implies that different mechanisms are used for spatial polarization of membrane proteins than for polarization of cytoplasmic elements such as organelles, messenger RNA and structural proteins — including components of the initial segment itself. □

Peter J. Hollenbeck and Gordon Ruthel are in the Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907, USA. e-mail: phollenb@purdue.edu

1. Brown, D. & Stow, J. L. *Physiol. Rev.* **76**, 245–297 (1996).
2. Winckler, B., Forcher, P. & Mellman, I. *Nature* **397**, 698–701 (1999).
3. Winckler, B. & Poo, M.-M. *Nature* **379**, 114 (1996).
4. Bennet, V. & Gilligan, D. M. *Annu. Rev. Cell Biol.* **9**, 27–66 (1993).
5. Davis, J. Q. & Bennet, V. J. *Biol. Chem.* **269**, 27163–27166 (1994).
6. Davis, J. Q., Lambert, S. & Bennet, V. J. *Cell Biol.* **135**, 1355–1367 (1996).
7. Simons, K. & Ikonen, E. *Nature* **387**, 569–572 (1997).
8. Craig, A. M. & Banker, G. *Annu. Rev. Neurosci.* **17**, 267–310 (1994).

RNA splicing

Running rings around RNA

Angus I. Lamond

Twenty years ago it was discovered that many people suffering from autoimmune diseases produce autoantibodies against Sm proteins — a family of proteins that bind to nuclear RNA. Subsequent work showed that the Sm proteins are involved in removing introns (regions not encoding the protein product) from messenger RNA precursors. Now, a study by Kambach *et al.*¹ in *Cell* reports the crystal structures of two Sm protein complexes, and offers new insights into the architecture of the RNA splicing machinery.

Almost all genes in higher eukaryotes contain one or more introns that must be removed from pre-mRNA transcripts to produce functional mRNA. This RNA splicing takes place in the nucleus, and is catalysed by a large RNA–protein complex known as the spliceosome. In both yeast and mammals, the spliceosome contains conserved RNA–protein subunits, the U1, U2, U4/U6 and U5 small nuclear ribonucleoprotein particles (snRNPs, pronounced “snurps”).

These snRNP subunits, together with additional protein-splicing factors, bind to conserved RNA sequence motifs at either end of the intron to form an active spliceosome.

Each snRNP subunit comprises one or two small nuclear RNAs (snRNAs), and a group of proteins that includes the Sm proteins². Although some of these proteins are specifically associated with just one type of snRNP, the Sm proteins bind to a conserved RNA motif — the Sm RNA motif — that is present in the U1, U2, U4 and U5 snRNAs. This core of Sm proteins, in turn, consists of seven polypeptides called B/B', D₁, D₂, D₃, E, F and G (Fig. 1). (The B/B' proteins are alternatively spliced isoforms of the same gene.)

Interest in snRNPs has been stimulated by evidence³ that an inherited human motor-neuron degenerative disease, spinal muscular atrophy, may result from failure of the Sm protein complex to assemble. To assemble a snRNP, the U snRNAs are first transcribed in the nucleus by RNA polymerase II and, like mRNAs, they receive an

N7-methyl-guanosine 5' cap. They are then exported to the cytoplasm where the Sm proteins bind and trigger hypermethylation of the 5' cap structure. Finally, the snRNP particle is reimported into the nucleus.

Biochemical studies⁴ have shown that, in the absence of U snRNA, the seven Sm polypeptides can form three stable sub-complexes — D₁ and D₂; E, F and G; and D₃ and B/B'. These sub-complexes do not bind U snRNAs alone, but D₁D₂ and EFG can form a stable partial complex with the Sm RNA motif and, if D₃B/B' is then added, a complete core snRNP complex is produced. Kambach *et al.*¹ have now solved the crystal structures of two of these Sm sub-complexes, D₁D₂ and D₃B, at resolutions of 2.5 and 2 Å, respectively and, from these data, have come up with a model for the overall organization of the Sm core domain and how it may bind RNA.

Sequence comparisons^{5–7} from different species have previously shown that the Sm proteins belong to a conserved family, typified by a two-part amino-acid motif (the Sm motif). The two conserved segments, called Sm1 and Sm2, are separated by a linker region of variable length, and they form a stable domain that mediates interactions between the Sm proteins. Kambach *et al.* now show that the D₃ and B proteins have a common fold, which contains an amino-ter-

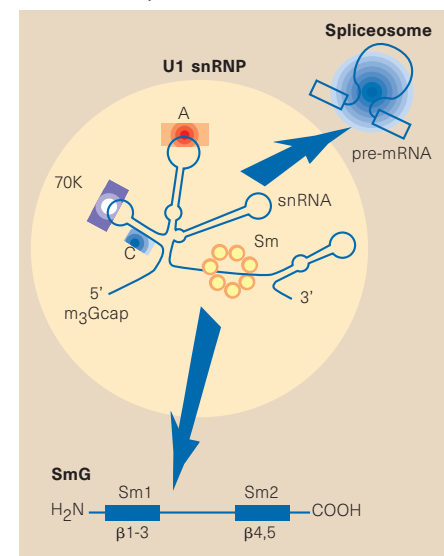


Figure 1 Organization of the spliceosome. In yeast and mammals the spliceosome contains small nuclear ribonucleoprotein particles (snRNPs) U1, U2, U4/U6 and U5. The snRNPs contain one or two small nuclear RNAs (snRNA), along with a group of proteins that includes the Sm proteins. Each Sm protein binds to a conserved region of the snRNA called the Sm RNA motif, and is made up of a core of polypeptides. In the absence of RNA, these polypeptides can organize themselves into three sub-complexes, two of which have now been crystallized by Kambach *et al.*¹. Each Sm protein contains two conserved domains, Sm1 and Sm2, joined by a variable linker, and the strands of a five-stranded β -sheet span these two domains.

minal α -helix followed by a strongly bent, five-stranded, antiparallel β -sheet. The first three β -strands are part of the conserved Sm1 motif, whereas the fourth and fifth strands belong to the Sm2 motif.

The authors have determined the structures of the bent β -sheets for D₁, D₂, D₃ and B, and find that their Sm1 and Sm2 motifs closely superimpose. This indicates that the conserved amino acids of the Sm motif form a specific structural domain in Sm proteins. By comparing the interaction surfaces formed between D₁ and D₂, and D₃ and B, Kambach *et al.* also found that the overall subunit geometry between the two subcomplexes is well conserved. So, the fourth β -strands of D₂ and B interact with the adjacent fifth β -strands of D₁ and D₃, respectively and, in both cases, the interaction is stabilized by main-chain hydrogen bonding.

This conservation of the subunit interfaces between Sm proteins led Kambach *et al.* to predict the structures of the E, F and G proteins. Using these predictions, coupled with existing biochemical information about how the Sm proteins interact, they have produced a very plausible model for the overall structure of the core Sm complex (Fig. 2). The authors predict that seven Sm proteins form a closed ring, with the fourth β -strand of each protein interacting with the fifth β -strand of the adjacent protein in the ring. The proteins are arranged in the order D₂, D₁, B, D₃, G, E and F.

An interesting feature of the model is that the ring has a highly basic central hole into which polypeptide loops from each Sm protein protrude. This suggests that the Sm RNA may pass through the hole. In this way, the RNA can be stabilized both by the positive charges on the ring, and by specific interactions with amino-acid side chains in the pro-

truding loops. Although the central hole is too small for double-stranded RNA, it could accommodate the single-stranded Sm RNA sequence. And it is tempting to speculate that this core Sm protein structure represents an ancestral form of RNA-binding domain — the functional and structural conservation of Sm proteins, from yeast through to mammals, is consistent with this idea.

We now await confirmation of the model, in the form of a co-crystal of Sm proteins bound to RNA. The world of RNA-protein structure is relatively unexplored territory, and it will be interesting also to see whether the seven-membered ring emerges as a

common building block of RNA-protein complexes in other biological systems. □

Angus I. Lamond is in the Department of Biochemistry, University of Dundee, MSI/WTB Complex, Dow Street, Dundee DD1 5EH, UK. e-mail: a.i.lamond@dundee.ac.uk

1. Kambach, C. *et al.* *Cell* **96**, 375–387 (1999).
2. Lührmann, R., Kastner, B. & Bach, M. *Biochim. Biophys. Acta* **1087**, 265–292 (1990).
3. Fischer, U., Liu, Q. & Dreyfuss, G. *Cell* **90**, 1023–1029 (1997).
4. Raker, V. A., Plessel, G. & Lührmann, R. *EMBO J.* **15**, 2256–2269 (1996).
5. Cooper, M., Johnston, L. H. & Beggs, J. D. *EMBO J.* **14**, 2066–2075 (1995).
6. Hermann, H. *et al.* *EMBO J.* **14**, 2076–2088 (1995).
7. Séraphin, B. *EMBO J.* **14**, 2089–2098 (1995).

Materials science

Measures of crystal vacancies

Robert W. Cahn

It is 73 years since Frenkel¹ suggested that some atoms in a crystal will spontaneously leave their lattice sites so that a proportion of vacant lattice sites, or vacancies, remain in equilibrium. That notion, for the first time, explained how matter can be transported in a crystalline solid, because neighbouring atoms can drop into a vacancy, which thus itself migrates. Ever since then, crystal vacancies have been central to solid-state physics, metallurgy and, eventually, materials science. A methodological innovation in the study of such vacancies now comes from Schaefer *et al.*², writing in *Physical Review Letters*. They show how the time variation of vacancy concentration in a crystal can be experimentally determined at varying temperatures. There is currently great interest in the effects of large vacancy concentrations on the mechanical properties of the alloys FeAl and NiAl, and the new technique offers a powerful method of investigating these phenomena.

The first measurements of point-defect concentrations go back to 1943, when Stern (a German refugee in the United States) developed an ultraprecise liquid column with a slight density gradient: ionic crystals were floated in the column and came to rest at a location depending on their density. When such a crystal had been irradiated to generate 'F-centres' (a topic of intense interest at that time), a change of density of one part in 10⁴ could be accurately measured, and so the identity of F-centres as associated with vacancies at last became clear. This work was not published³ until 1949; in the meantime, Stern had gained a Nobel prize.

The method was not applicable to metals, if only because their density was too high for any available suspension liquids. In 1958, Feder and Nowick⁴ invented a method that has been developed further ever since: "Their concept involved a comparison of the macroscopic thermal expansion, as measured by the overall volume change of

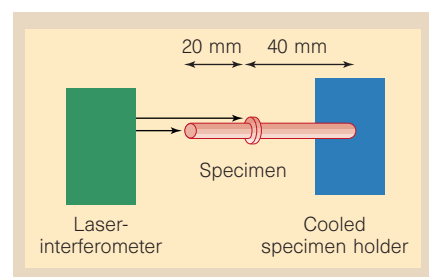


Figure 1 The experimental approach employed by Schaefer *et al.*². The specimen is suspended together with a furnace in a vacuum, and its length change (and therefore volume change) is measured by the two laser beams on the polished parallel planes on the specimen front and ledge. (From ref. 2.)

the sample, and the microscopic expansion, as measured by X-ray diffraction" (the quotation comes from Nowick's recent overview⁵ of "the golden age of crystal defects"). The difference in the two measurements cited is due to point defects.

The method was soon taken up and developed⁶, but a key step was subsequently taken by Feder, Nowick's collaborator, who replaced contact dilatometry by laser interferometry⁷. Later, another method of measuring vacancy concentration in equilibrium was invented: positron-annihilation spectrometry⁸. Positrons are injected into the crystal and many of them are annihilated by electrons near vacancy sites; the average lifetime of the positrons is related to vacancy concentration.

The 'volume comparison' approach was not, at the time, developed to study changes in the vacancy concentration to be measured. This could only be done by more indirect means, and Turnbull *et al.*⁹ indeed showed how electrical resistivity measurements could be used to map the quenching-out of nonequilibrium vacancies in an age-hardening aluminium alloy.

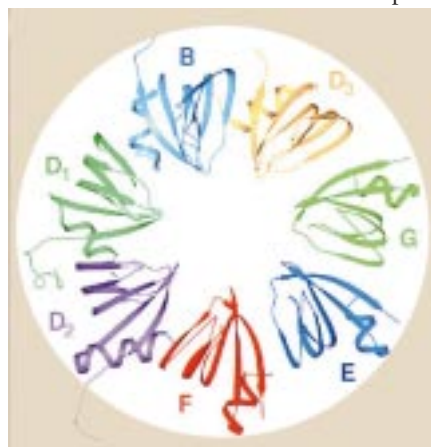


Figure 2 Proposed structure of the human Sm core complex¹. Based on the crystal structures of the D₁D₂ and D₃B Sm sub-complexes, Kambach *et al.*¹ propose that the seven core Sm proteins form a ring. The fourth β -strand of each protein interacts with the fifth β -strand of the next. Because the central hole that this creates is very basic, the authors suggest that Sm RNA could pass through its centre.

Schaefer *et al.*², who work in Stuttgart, have improved the volume comparison method by modifying Feder's approach. They used an 'incremental' Michelson laser interferometer to allow contact-free measurements as a function of time, up to high temperature and without contamination. Figure 1 shows the experimental setup: a ledge on the specimen is spark-machined *in situ* and polished to reflect the laser beam. In effect, a venerable instrument has been upgraded with modern electronics.

Schaefer and colleagues used the instrument to measure both the rate and the magnitude of change in vacancy concentration, following a sudden temperature change, in FeAl and NiAl intermetallics with the ordered B2 crystal structure. The time variation of the length change was modelled and accurately fitted to determine both formation energy and migration energy of the defect with good precision. In FeAl, two types of defect were detected: the minor constituent remains something of a mystery and may in fact be a cluster of point defects.

A group in Göttingen have simultaneously¹⁰ used a combination of dilatometry and positron annihilation to study equilibrium vacancies in FeAl, determining a vacancy concentration at stoichiometry as high as 3.5% at the melting-point, consistent with the low formation energy measured by the Stuttgart team. This vacancy concentration is of particular interest because of the growing attention being paid to the relationship between vacancy concentration and mechanical strength in compounds with the B2 (CsCl-type) crystal structure. This idea was put on the map in a study¹¹ of hardness in relation to vacancy concentrations calculated from thermodynamic data. The new technique² makes it possible to measure the relevant vacancy concentrations directly, and also the rate at which they change when the temperature is suddenly changed.

George and Baker¹² have surveyed the literature of the past few years on this form of vacancy-hardening, and go on to explain how such hardening, varying sharply with temperature because of changes in vacancy concentration, can explain the anomalous yield effect in FeAl: the yield strength decreases at first with rising temperature but then inverts and peaks at about 600 °C, before dropping rapidly with further temperature increase. The rapid drop is ascribed to dislocation-creep drastically accelerated by a high vacancy concentration. To confirm their model, they used an instrument designed for welding researchers to simulate rapid temperature changes in a weld zone; specimens heated in this device were mechanically tested at various times after the fast heating-up. Now, it appears, the approach of Schaefer *et al.* could be used to directly measure vacancy concentrations after various heat-treatments, and in this

way could help check the George-Baker theory of anomalous yield stress in FeAl. □

Robert W. Cahn is in the Department of Materials Science and Metallurgy, University of Cambridge, Pembroke Street, Cambridge CB2 3QZ, UK.
e-mail: rwc12@cam.ac.uk

1. Frenkel, J. *Z. Physik* **35**, 652–669 (1926).
2. Schaefer, H.-E., Frenner, K. & Würschum, R. *Phys. Rev. Lett.* **82**, 948–951 (1999).
3. Estermann, J., Leivo, W. J. & Stern, O. *Phys. Rev.* **75**, 627–633 (1949).

4. Feder, R. & Nowick, A. S. *Phys. Rev.* **109**, 1959–1963 (1958).
5. Nowick, A. S. *Annu. Rev. Mater. Sci.* **26**, 1–19 (1996).
6. Simmons, R. O. & Balluffi, R. W. *Phys. Rev.* **119**, 600–605 (1960); **125**, 862–872 (1962).
7. Feder, R. & Charbneau, H. *Phys. Rev.* **149**, 464–471 (1966).
8. Wollenberger, H. J. in *Physical Metallurgy*, 4th edn (eds Cahn, R. W. & Haasen, P.) 1633 (North-Holland, Amsterdam, 1996).
9. Turnbull, D., Rosenbaum, H. S. & Treafits, H. N. *Acta Metall.* **8**, 277–295 (1960).
10. Wolff, J. *et al. Intermetallics* (in the press).
11. Chang, Y. A. *et al. Intermetallics* **1**, 107–115 (1993).
12. George, E. P. & Baker, I. *Intermetallics* **6**, 759–763 (1998).

Climate change

The question of significance

A. Barrie Pittock

Projections of rising greenhouse gas concentrations in the atmosphere, their climatic consequences, and their social and environmental effects are becoming increasingly sophisticated¹. Nevertheless, these studies invariably highlight the need for further improvements in the data and models, and in their analysis. Such is the case with the paper by Hulme and colleagues on page 688 of this issue², which nonetheless constitutes a considerable advance.

The authors have combined results from simulations of changing climate and its environmental effects to look at Europe in the year 2050. They ask whether the impacts of human-induced climate change, through greenhouse gas emissions, can be distinguished from those due to natural climate variability. The two indicators they use are river flow and wheat yield. Both are affected by temperature and rainfall.

Climate change modelling has come a long way since the early days. In the 1970s, most climate change scenarios and impact analyses were based on estimated global

average warmings and patterns derived from palaeoclimatic and other analogues³. Then came studies based on atmospheric general circulation models (GCMs) of coarse horizontal resolution, bounded at the surface by a simple 'slab-layer' ocean. These were run for changed atmospheric composition, usually for equilibrium doubled CO₂ concentrations. Changes at times earlier than that of doubled CO₂ were obtained by linear scaling. Due to computing limitations, the equilibrated control (present climate) and doubled CO₂ climate simulations were usually run only over a period of one or two decades⁴.

During the late 1980s, with the aid of more powerful computers, climate modellers progressed to simulations with fully coupled ocean-atmosphere GCMs which explicitly model ocean dynamics and oceanic heat uptake⁵. These enable simulations with continuous variations of greenhouse gas concentrations from the present or earlier (preindustrial) conditions, following some emission scenario path to 2100 or



Figure 1 Harvesting wheat in Denmark. Wheat growth is highly sensitive to temperature and rainfall, and thus to climate change. In their simulations to the year 2050, Hulme *et al.*² use wheat yield and river flow in several European countries as impact indicators. The question they tackle is whether the possible impacts of human-induced climate change can be distinguished from those stemming from natural variability in climate.

beyond. Even so, many impact studies have relied on analyses of only a few decades of output from the longer control and enhanced greenhouse simulations. A severe problem with these studies is that of 'natural' multi-decadal variations in the simulated control climates, and indeed in the simulated changing climates. Such variations may dominate the analyses.

Hulme *et al.*² help to put this problem in perspective by attempting to quantify the impacts of human-induced climate change relative to those of natural climate variability. They use a multi-century simulation of the control climate, run by the Hadley Centre in Britain, to extract seven more-or-less independent '30-year climates', and thus to provide an estimate of the simulated natural variability in climate. They also carried out multiple simulations, or 'ensemble runs', for changed climate under projections of 0.5% and 1% annual increases in equivalent CO₂ concentrations (for comparison, the observed rate of increase over the past decade, while fluctuating, is nearer to 0.5%).

Hulme *et al.* used their estimates of monthly climate change from ensemble means from four simulations, interpolated to a 0.5° latitude/longitude grid over Europe, to estimate changes in river flow and wheat yield using two climate-impact models. They then compared the simulated impacts due to human-induced climate change with the maximum simulated impacts due to the estimated natural variability. For around 2050, they find that climate change scenarios show systematic increases in river flow in northern Europe, and decreases in southern Europe, both being statistically significant relative to natural variability. But central and western Europe show smaller and non-significant changes. The spatial patterns in the climate change cases are far more consistent than in the climate variability cases.

Wheat yield in Europe is highly sensitive to natural variability in temperature and rainfall. In these simulations, climate change alone by 2050 results in significant increases in yield under both scenarios only in three northern European countries. But adding estimated effects of higher CO₂ concentrations on plant growth increases the yields considerably, with yield increases now being significant in all of the ten countries included in the study. Whether such increases due to higher CO₂ concentration would actually be realized is uncertain, because of problems in extrapolating from idealized laboratory and modelling results to field conditions⁶.

The clear message of this work is that greater efforts are needed to take account of the 'noise' of natural climate variability when considering the 'signal' of climate change. In particular, care must be taken in interpreting the results of impact assessments, especially those which use climate simulations over short periods only.

Much more remains to be done. For example, the use of simulated monthly climate data by Hulme *et al.*², even when interpolated to daily data by statistical methods, leaves open the possibility of systematic changes in day-to-day variability and extremes due to climate change^{7,8}. These could well dominate the effects of climate change on both river flow and wheat yield. Taking them into account will require the analysis of daily output from the climate models, especially at fine spatial resolution.

More importantly, in understanding the message from Hulme *et al.*, we need to distinguish between confidence levels used for the detectability of simulated or real impacts due to climate change, and those needed to decide if the information about possible future impacts is useful for decision-makers. Detectability, used as evidence of climate change, arguably requires a high level of confidence, say at the 95% or 99% level. On the other hand, if we take it as a given that climate change is occurring, and we are seeking useful advice as to what the consequences might be, a lower level of confidence may suffice.

Studies of the length of a data set necessary to detect a real change in the mean against a background of variability, for example for detection of trends in total ozone⁹ or river flow¹⁰, indicate that even if the change is real, it may need to have happened (and thus will have had impacts) for decades before it can be shown statistically to have occurred. In cases such as estimation of flood frequency and of the occurrence of moisture stress in crops, what is needed to identify future 'dangerous' levels of climate change, or to plan adaptation measures, is an estimate of changes in the probability of such effects before they happen. This is consistent with the 'precautionary principle' embodied in the UN Framework Convention for Climate Change. Given the continuing uncertainties in estimating future climate changes and impacts, a risk assessment approach is essential¹¹. □

A. Barrie Pittock is at CSIRO Atmospheric Research, PMB 1, Aspendale 3195, Australia.
e-mail: barrie.pittock@dar.csiro.au

- Pittock, A. B. in *Modelling Change in Environmental Systems* (eds Jakeman, A. J., Beck, M. B. & McAleer, M. J.) Ch. 20 (Wiley, Chichester, 1993).
- Hulme, M. *et al.* *Nature* **397**, 688–691 (1999).
- Pittock, A. B. & Salinger, M. J. *Clim. Change* **4**, 23–40 (1982).
- Mitchell, J. F. B., Manabe, S., Melieshko, V. & Tokioka, T. in *Climate Change: The IPCC Scientific Assessment* (eds Houghton, J. T., Jenkins, G. J. & Ephraums, J. J.) 131–172 (Cambridge Univ. Press, 1990).
- Kattenberg, A. *et al.* in *Climate Change 1995: The Science of Climate Change* (eds Houghton, J. T. *et al.*) 285–357 (Cambridge Univ. Press, 1996).
- Pittock, A. B. *Environment* **37** (9), 25–30 (1995).
- Hennessy, K. J., Gregory, J. M. & Mitchell, J. F. B. *Clim. Dyn.* **13**, 667–680 (1997).
- Hennessy, K. J. & Pittock, A. B. *Int. J. Climatol.* **15**, 591–612 (1995).
- Pittock, A. B. *Pure Appl. Geophys.* **118**, 643–661 (1980).
- Chiew, F. H. S. & McMahon, T. A. *Int. J. Climatol.* **13**, 643–653 (1993).
- Pittock, A. B. & Jones, R. N. *Environ. Monitor. Assess. J.* (in the press).

Daedalus

Pinpoint therapy

The medical hypodermic syringe is a shotgun. It injects a gram or so of drug solution, which then diffuses wastefully through the whole body. Daedalus is now refining it. His 'microdermic syringe' has a needle a micrometre or so across, smaller than a typical cell diameter. It will nose its way between the cells of the body without hurting or doing damage. It will pass harmlessly through organs and blood-vessels, and even through the microporosity of bone. It will deliver a drug to its precise target, however deep.

The ideal microdermic needle would simply be a giant multi-layered carbon nanotube — far stronger and stiffer than steel. Carbon technology is advancing so fast that large nanotubes with thick enough walls should soon be available. For microdermic use, a nanotube should be treated with fluorine. This would coat its outer surface with slippery graphite fluoride, minimizing the friction of its insertion. Of course, such a long fine needle could not just be pushed into the skin; it would buckle. It must be driven from the skin surface itself, perhaps by tiny traction rollers. Daedalus even hopes that it could be drawn in by repeated ultrasonic pulses sent down it, and reflected at its tip as tensile pulls (a scheme he has proposed for driving nails). If that tip were angled, the pulses would tend to draw it in along a gentle curve. The needle could then be steered into the body by rotating it from the surface, so as to vary the direction of that curve. Much of the arcane know-how of the oil-drilling industry might be scaled down for the job. Some sort of imaging system would be needed to guide the needle to its destination. The ultrasonic pulses radiating from its tip could locate it in sonar-fashion.

An aqueous solution would take ages to flow down such a fine bore. So Daedalus will dissolve the drug in a supercritical solvent, with its low viscosity and high pressure. In supercritical carbon dioxide under 73 atmospheres of driving pressure, it will traverse the needle far faster. At the site of delivery, the few micrograms of carbon dioxide will dissolve harmlessly in the body's fluids. The drug will act exactly where it is wanted, and nowhere else; side-effects will vanish. The needle could be even be left in place. A tiny sub-dermal pressure-capsule could leak the drug continuously into the target till the trouble was fully cleared, by maybe a millionth of the usual dose. The drug companies would be furious.

David Jones

Growing season extended in Europe

Changes in phenology (seasonal plant and animal activity driven by environmental factors) from year to year may be a sensitive and easily observable indicator of changes in the biosphere. We have analysed data from more than 30 years of observation in Europe, and found that spring events, such as leaf unfolding, have advanced by 6 days, whereas autumn events, such as leaf colouring, have been delayed by 4.8 days. This means that the average annual growing season has lengthened by 10.8 days since the early 1960s. These shifts can be attributed to changes in air temperature.

Forests in cool and temperate zones are adapted to the seasonal cycle, having a dormancy period during winter that is triggered mainly by seasonal variation in temperature and light. Seasonal phases can also be influenced by soil, water supply and biotic factors, including genes^{1,2}. Plants can therefore be used as biological indicators of changing environmental conditions, with springtime phases being particularly sensitive to temperature.

We have been analysing³ observational data from the International Phenological Gardens (IPG), a Europe-wide network covering a large latitudinal (from Scandinavia at 69° N to Macedonia at 42° N) and longitudinal (from Ireland at 10° W to Finland at 27° E) range that has genetically identical clones of trees and shrubs (Fig. 1). The dates of observed phases, such as leaf unfolding, May shoot, flowering, leaf colouring and leaf fall, were recorded from 1959 to 1993.

We set out to apply a unified phenological model to the IPG network in order to simulate the consequences of climatic change. According to the model¹, more than 70% of the interannual variability in

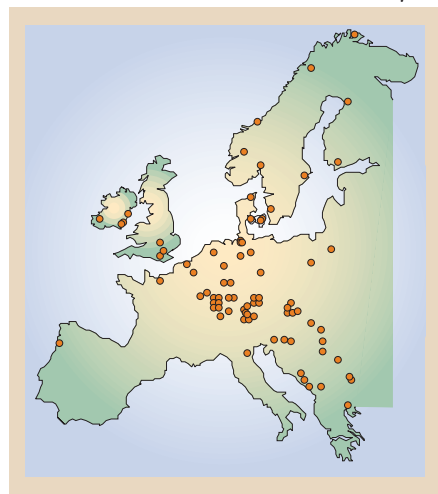


Figure 1 Sites of the International Phenological Gardens (IPG).

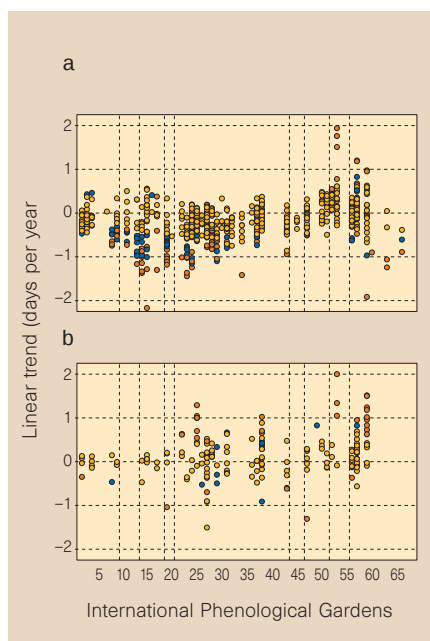


Figure 2 Linear trends of phenological phases. Data are for long observational series (20 years or more) during the 1959–93 period. Orange dots, significant at the 1% level; blue dots, significant at the 5% level; and yellow dots, not significant at the 5% level; *F*-test. a, Spring phases (including leaf unfolding, May shoot, flowering); b, autumn phases (including leaf colouring, leaf fall).

bud-break can be explained by daily temperatures; mean absolute errors in estimating phenological data ranged between 4 and 8 days. The model showed that the onset of spring events is sensitive to climatic change and will advance by up to six days per 1 °C increase in winter air temperature depending on the species, a figure similar to previous results^{4,5}.

We also looked for trends that might reflect global warming, so, for all phenological records covering at least 20 years, we calculated linear trends and their significance according to *F*-test, trend/noise ratio and the Mann–Kendall trend test. Only a quarter of all trends are significant. Spring phases generally show a tendency for negative trends (Fig. 2a), indicating that spring is starting earlier in later years. Some of the Balkan stations, in contrast, show higher positive trends, indicating that spring is delayed, but this is probably the result of regional patterns of climate change. A positive (but less consistent) trend for the autumn phases (Fig. 2b) indicates that the beginning of autumn is gradually delayed over the years.

Averaging all 616 springtime data series reveals a negative trend of -0.20 days per year, equivalent to an advance of 6 days over 30 years. Analysis of the 178 autumn data

series reveals a mean positive trend of $+0.16$ days per year, which would mean a delay of 4.8 days over 30 years. The growing season over this period has therefore lengthened by 10.8 days on average.

Other factors that could result in similar trends can be excluded. Only a few of the IPGs are situated in city areas where the urban heat island could influence these trends. In addition, the trends of spring phases are not an artefact of plant ageing because the increasing age of trees would be expected to delay the start of spring.

Our Europe-wide results are supported by regional studies, including a shift towards earlier flowering of the locust tree *Robinia pseudoacacia* in Hungary⁶ and the advance of spring and delay of autumn revealed by analysis of thermal seasons in Germany⁷.

The apparent increase of 10.8 days in the mean annual growing season in Europe is in accord with an analysis of satellite data from 1981 to 1991, which estimates a global advance of 8 ± 3 days for the beginning of the growing season and a prolongation by 4 ± 2 days of the declining phase⁸. Our data also confirm a reported⁹ advance in the seasonal cycle of about 7 days that was inferred from CO₂ data taken between the 1960s and early 1990s, with most of the effect occurring after 1980.

Our phenological modelling indicates that the changes are caused by a temperature increase, which is one of the effects of global warming. The lengthening of the growing season is likely to contribute to increased biomass formation, which is part of a global increase in biospheric activity^{8,9}. This increase has resulted in accelerated tree growth across Europe, which has previously been attributed¹⁰ to fertilization by nitrogen compounds and CO₂, but the effects of lengthening the growing season have yet to be quantified.

Annette Menzel, Peter Fabian

Department of Bioclimatology and Pollution Research, University of Munich, Faculty of Forest Science, D-85354 Freising, Germany
e-mail: menzel@met.forst.uni-muenchen.de

- Defila, C. Thesis, Univ. Zürich (1991).
- Schnelle, F. *Pflanzen-Phänologie* (Geest & Portig, Leipzig, 1955).
- Menzel, A. *Forstliche Forschungsberichte* 164 (Münich, 1997).
- Park-Ono, H. S., Kawamura, T. & Yoshino, M. in *Proc. 13th Int. Congr. Biometeorology, Sept 12–18, 1993* (eds Maarouf, A. R. & Barthakur, N. N.) 207–220 (Calgary, 1994).
- Kramer, K. *Plant Cell Environ.* **17**, 367–377 (1994).
- Walkovszky, A. *Int. J. Biometeorol.* **41**, 155–160 (1998).
- Rapp, J. & Schönwiese, C.-D. *Meteorol. Z.* **3**, 91–94 (1994).
- Myneni, R. B., Keeling, C. D., Tucker, C. J., Asrar, G. & Nemani, R. R. *Nature* **386**, 698–702 (1997).
- Keeling, C. D., Chin, F. J. S. & Whorf, T. P. *Nature* **382**, 146–149 (1996).
- Spiecker, H., Mielikäinen, K., Köhl, M. & Skovsgaard, J. P. (eds) *Growth Trends in European Forests: Studies from 12 Countries* (Springer, Berlin, 1996).

Breathing through skin in a newborn mammal

The exchange of oxygen and carbon dioxide through the skin does not occur in most mammals because they have high metabolic rates and diffusion through the skin is poor. But we have found that in the Julia Creek dunnart (*Sminthopsis douglasi*), a marsupial mouse with one of the smallest newborns of any mammal, gas exchange through the skin is the predominant form of O₂ and CO₂ transfer in the first days after birth.

These small newborns have lower O₂ consumption than non-marsupial species¹, and thermoregulation is not a problem because they develop in the thermoneutral environment of the maternal pouch. Further, because they are delivered at a very early stage of development², the skin, which is hairless and rich in blood, provides a much smaller barrier to gas diffusion than in most newborn mammals.

A 70-kg human athlete can reach a level of maximal oxygen consumption (V_{O_2}) of up to 65 ml kg⁻¹ min⁻¹ (4,550 ml min⁻¹). Such high values are possible because of a very large surface area for gas exchange in the lungs, estimated at about 70 m² (ref. 3). In an adult human, the body surface could therefore never replace the lungs as a site of gas exchange because its total area is only about 2% of the minimal surface required. If the lungs were designed optimally⁴, the ratio between the pulmonary gas-exchange surface and O₂ flux would be the minimum necessary for gas exchange. In humans, this value would be 70 m² per 4,550 ml min⁻¹, or about 150 cm² per ml O₂ min⁻¹.

The Julia Creek dunnart is a small dasyurid marsupial from Australia⁵. After a



Figure 1 Pouch young of *Sminthopsis douglasi* at one day old. The lungs are visible as air sacs on each side of the heart. Despite the presence of ribs, no respiratory thoracic movements can be detected. Scale bar, 1 mm.

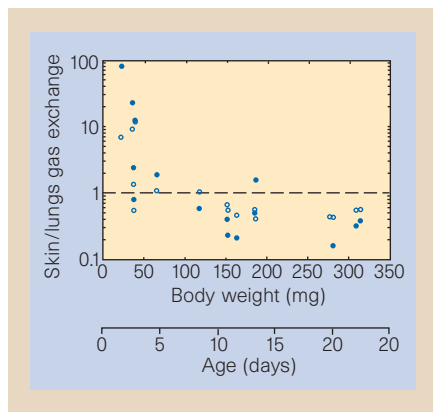


Figure 2 Ratio between skin and pulmonary gas exchange in the pouch young of *S. douglasi*. Values refer to oxygen consumption (V_{O_2} , filled symbols) and carbon dioxide production (V_{CO_2} , open symbols), measured at an ambient temperature of 36 °C. The broken line indicates equal contributions of lungs and skin to total gas exchange.

gestation lasting about 13 days, the newborn is about 4 mm long and weighs around 17 mg, so it is one of the smallest newborn mammals known². At birth, the skeleton is entirely cartilaginous and the internal organs are visible through the transparent skin. The lungs are represented by a small number of approximately spherical air sacs (Fig. 1). In the newborn dunnart and in individuals up to 21 days old, V_{O_2} averaged 18 ml kg⁻¹ min⁻¹, and the body surface area is estimated at about 7 mm². The ratio between body surface area and V_{O_2} is therefore about 220 cm² per ml O₂ min⁻¹, above the minimum value necessary for the skin to be an important site of gas exchange.

We measured separately, but simultaneously, the gaseous metabolism of lungs and skin in 22 pouch young from five litters during the first three weeks after birth. We sealed a mask made from a short length of polyethylene tube to the face of the animal, covering both mouth and nostrils, using a removable dental polyether material. The low mobility of the newborns ensured that the seal was maintained. The tube passed through a thin rubber stopper placed in the centre of a moist cylindrical chamber of volume 0.5–2 ml, completely separating the chamber into two compartments, one containing the animal and the other communicating with the airways.

A small quantity of air was injected into one compartment, and the absence of pressure transmission to the other indicated complete separation. The chamber was maintained at pouch temperature (36 °C) by a water bath. After temperature and humidity equilibration, the compartments were sealed for 5–15 min, depending on the animal's age. The compartments were then flushed with a constant airflow of 20 ml min⁻¹ and the gas was forced

through a drying column before O₂ and CO₂ concentration were determined.

V_{O_2} and V_{CO_2} (CO₂ production) were calculated from the time integral of the gas concentration curves⁶, multiplied by the flow and the time the chamber was sealed. At all ages, the skin's contribution to gas exchange was very marked. In the youngest animals, with body weight below 100 mg, gas exchange through the skin exceeded that through the lungs (Fig. 2). In the oldest individuals studied, which were 20 to 21 days of age with a body weight of about 290 mg, skin exchange was about one-third that of the lungs.

In animals two to three days old, spontaneous body movements did not appreciably expand the air sacs, and resulted in minimal changes in lung volume. These were calculated from the displacement of a drop of soapy solution in a microtube directly connected to the polyethylene tube sealed to the face of the animal, magnified by a microscope and displayed on a television screen by a video camera. These observations suggest that, in the Julia Creek dunnart during the early postnatal phases, pulmonary convection is highly inefficient. Sustaining oxygen demands through the skin allows these very small animals to be born before the respiratory apparatus is fully functional.

J. P. Mortola*, P. B. Frappell†, P. A. Woolley†

*Department of Physiology, McGill University, Montreal, Quebec, H3G 1Y6, Canada

†Department of Zoology, La Trobe University, Melbourne, Victoria 3083, Australia
e-mail: jacopo@physio.mcgill.ca

1. Baudinette, R. V., Gannon, B. J., Ryall, R. G. & Frappell, P. B. *Respir. Physiol.* **72**, 219–228 (1988).
2. Tyndale-Biscoe, H. & Renfree, M. *Reproductive Physiology of Marsupials* (Cambridge Univ. Press, 1987).
3. Weibel, E. R. *Morphometry of the Human Lung* (Springer, Berlin, 1963).
4. Weibel, E. R. *The Pathway for Oxygen: Structure and Function in the Mammalian Respiratory System* (Harvard Univ. Press, Cambridge, Massachusetts, 1984).
5. Woolley, P. A. in *The Mammals of Australia* (ed. Strahan, R.) 134–135 (Reed, Chatswood, New South Wales, 1995).
6. Frappell, P. B., Blevin, H. A. & Baudinette, R. V. *J. Theor. Biol.* **138**, 479–494 (1989).

Formation of ice XII at different conditions

The structural versatility of solid water has been reinforced by the identification of a metastable, crystalline phase called ice XII (ref. 1). It was obtained by cooling liquid water to 260 K at a pressure of 0.55 GPa. We have found ice XII in a completely different region of water's phase diagram and show that it can be formed under different conditions from those previously reported¹.

We produced high-density amorphous ice (HDA) as described previously^{2,3} and

observed crystalline impurities⁴ within the amorphous samples⁵. These impurities are metastable at a temperature of 77 K and ambient pressure. Since the discovery of ice XII, we have identified this crystalline contaminant. We used Rietveld refinement to confirm that the crystalline phase has a tetragonal unit cell with cell parameters $a = 8.276 \pm 0.004$ Å and $c = 4.027 \pm 0.004$ Å, leading to a density, ρ , of 1.30 g cm^{-3} (H_2O) at 127 K and ambient pressure (ρ for HDA is 1.17 g cm^{-3}).

A neutron powder-diffraction pattern of a fully deuterated sample is shown in Fig. 1a. The highest-symmetry space group meeting the reflection conditions is $I42d$. The calculated fractional coordinates and the intra- and intermolecular distances (see Supplementary Information) are in good agreement with previous data¹. The structure of the high-density crystalline phase therefore corresponds to ice XII. In all our samples, the contaminating crystalline phases can be indexed to ice XII, with no other crystalline modifications being formed. Ice XII is metastable at ambient pressure, and transforms at temperatures above 135 K to cubic ice (Fig. 1a, inset).

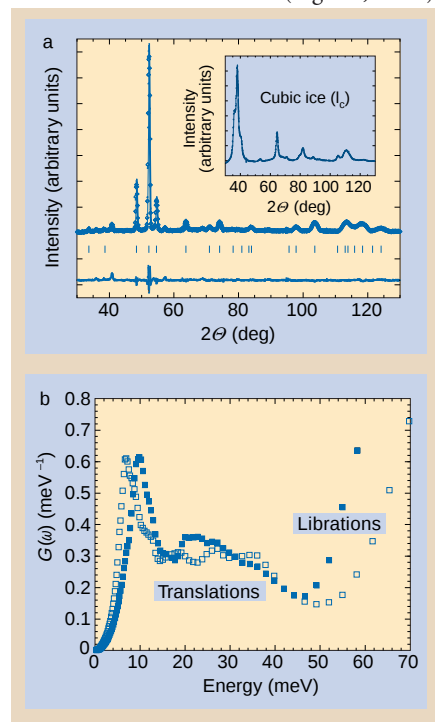


Figure 1 Elastic and inelastic neutron data. **a**, Measured and calculated neutron-diffraction patterns and the resulting difference profile of ice XI at 127 K and ambient pressure. The data were collected at the high-flux beamline D2O at the Institut Laue-Langevin, at a wavelength of 2.41 Å. Ticks indicate reflection positions. The inset shows the cubic phase (I_c) obtained after the transformation from ice XI to I_c . The shoulders on the I_c reflections are due to stacking faults¹⁰. **b**, Frequency distribution $G(\omega)$ of protonated ice XII (at 127 K, filled symbols) and I_h (at 240 K, open symbols) measured at ambient pressure.

This conversion takes about 35 min at 145 K.

To support our structural analysis, we determined the frequency distribution $G(\omega)$ of the external mode spectrum of ice XII in comparison with common hexagonal ice (I_h), measured on fully protonated samples by inelastic neutron scattering (Fig. 1b). Although the overall shape of $G(\omega)$ for ice XII is similar to those of other high-density ice phases, the details of the inelastic response are different⁶. $G(\omega)$ reflects the specific hydrogen-bond distortions that are unique to ice XII because it has seven- and eight-membered rings⁷. It therefore allows us to test hydrogen-bond models, such as the controversial idea⁶ that ice has two distinct hydrogen bonds.

We have tried to establish what it was in the production process that led to ice XII being produced. Its formation seems to be influenced neither by the shape nor by the material of the pressure cells. No distinct pressure dependence of the transition onsets for HDA and ice XII was observed, nor does the rate of compression favour one or the other of the two phases (rates from 1 GPa per 100 min up to 1 GPa per 100 s were applied to ensure isothermal and quasi-adiabatic conditions, in which heat is not lost or gained⁸). Samples containing pure D_2O (purity 99.9%, resistivity $1 \text{ M}\Omega \text{ cm}$) and H_2O (resistivity $18.2 \text{ M}\Omega \text{ cm}$) and isotopic mixtures give comparable results.

So we have been unsuccessful in identifying a route that would predictably allow us to separate the production of HDA from the production of ice XII. The relative amounts of HDA and ice XII are scattered more or less randomly, and both pure ice XII and pure HDA can be formed in some instances. Once HDA or ice XII has formed, it remains metastable up to a pressure of at least 2 GPa at 77 K.

These results indicate that water's phase diagram needs to be modified, particularly in the region that has been ascribed to HDA. It remains to be seen whether the region of metastability of ice XII previously reported¹ is continuously connected to the region observed here. An extensive study of the structural features, dynamics and kinetics accompanying transitions involving ice XII might also lead to a better understanding of HDA at a fundamental, microscopic level. For example, it might be possible to model the microscopic structure of HDA by putting disorder into the ice XII network, as has been done successfully for low-density amorphous ices using the hexagonal ice I network⁹.

Michael Koza*, Helmut Schober†, Albert Tölle*, Franz Fajara*, Thomas Hansen†

*Fachbereich Physik, Universität Dortmund, Otto-Hahn-Strasse 4, Dortmund 44221, Germany

e-mail: marek@e3.physik.uni-dortmund.de

†Institut Laue-Langevin, Avenue des Martyrs, BP 156, 38042 Grenoble, Cedex 9, France

1. Lobban, C., Finney, J. L. & Kuhs, W. F. *Nature* **391**, 268–270 (1998).
 2. Mishima, O., Calvert, L. D. & Whalley, E. *Nature* **310**, 393–395 (1984).
 3. Mishima, O., Calvert, L. D. & Whalley, E. *Nature* **314**, 76–78 (1985).
 4. Bizid, A., Bosio, L., Defrain, A. & Oumezzine, M. *J. Chem. Phys.* **87**, 2225–2230 (1987).
 5. Schober, H. *et al. Physica B* **241–243**, 897–902 (1998).
 6. Li, J. *J. Chem. Phys.* **105**, 6733–6755 (1996).
 7. O'Keefe, M. *Nature* **392**, 879 (1998).
 8. Mishima, O. *Nature* **384**, 546–549 (1996).
 9. Dore, J. C. in *Water Science Review* (ed. Franks, F.) 3–92 (Cambridge Univ. Press, 1985).
 10. Kuhs, W. F., Bliss, D. V. & Finney, J. L. *J. Physique* **48**, C-1631–C-1636 (1987).
- Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

The bio-logic of facial geometry

The final paragraph of the Letter by Perrett *et al.*¹, on the effects of sexual dimorphism on facial attractiveness, remarks that “preferences would encourage a youthful, neotenuous appearance in the species generally”. But this conflicts with the semantics of ‘masculine’ and ‘feminine’ they use to construct their argument. Brennan's caricatures of Ronald Reagan² also hint at this conflict. We have examined the logic of this argument and suggest that biological data might strengthen inferences made by Perrett *et al.*¹.

We characterize the argument as semantic because masculine and feminine are used to refer to three distinct aspects of faces: the extent to which they reflect the action of sex hormones during growth; how they are perceived by the subjects queried; and their position on the shape continuum defined by the 174 feature points measured¹.

We shall refer to these three aspects as biological, psychological and geometrical. The ‘surprising’³ nature of the observations rests on the unexamined identification of geometrical and biological or psychological sexuality: the idea that hormone-dependent sexual characteristics most notably^{4,5} advertise pathogen resistance⁶ (because the immunosuppressive consequences of testosterone and oestrogen production^{7,8} are Zahavian handicaps⁹) leads Perrett *et al.* to expect their experimental subjects to prefer masculinized male and feminized female faces¹. But this expectation evaporates if geometrical sexuality differs from biological sexuality for faces.

When their observations conflict with this expectation, Perrett *et al.* acknowledge that subjects' preferences “may reflect the effects of masculinity on perceived age”¹. That is, geometrical masculinity may be

correlated strongly with psychological and/or biological age so that geometrically feminized faces signal youthfulness more than biological femininity. If this were the case, then Perrett *et al.*'s observations would provide evidence for direct selection (for facial features signalling youthfulness), rather than against indirect selection (for secondary sexual characteristics signalling parasite resistance)¹⁰.

It is no surprise that geometrical extrapolation may signal characteristics other than those from which they are constructed. The choice of initial features can be almost arbitrary and can produce unintended signals. For example, Fig. 1 shows an extrapolation along a pongid to a hominid skull continuum comparing and contrasting gorilla, *Paranthropus*, Neanderthal and human skulls. For our present purposes, note the delicate 'feminine' jaw and round 'neotenuous' head. So the subjects studied by Perrett *et al.* might even be considered to be expressing a preference for 'hominized' over 'pongidized' faces.

To limit such spurious signals, we suggest that geometrical interpolation be used only to supplement biological data in constructing test faces: for example, to determine how subjects respond to age, averaged male and female faces could be prepared from sets of 15-year-old faces, from sets of 20-year-old faces, and so on. Ideally, averaged faces would represent a continuum of ages, but geometrical interpolation between closely spaced averages should be almost equivalent to a purely biological construction. Subjects could then be asked for their responses to these new facial subspaces.

The analogous experiment for biological sexuality is more difficult but still possible¹¹: sex-hormone titres should be measured during development and averages prepared from sets of faces grouped according to time-integrated hormone amounts. Again,

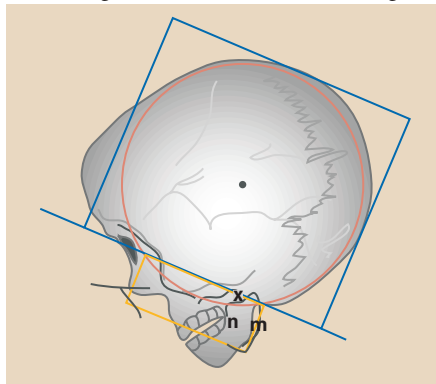


Figure 1 An extrapolation beyond human along a shape continuum defined by eight feature points on gorilla, *Paranthropus*, Neanderthal and human skulls. The subspace was constructed by hand (D.A.M.) using a procedure similar to, although considerably simpler than (that is, fewer skull features), the one followed by Perrett *et al.*¹. (Courtesy of Marilyn Pettit archives.)

to reduce the data collection, one could interpolate between some reasonable number of averaged faces. Furthermore, the resulting shape continua would, by construction, reflect biological sexuality rather than the geometrical version used by Perrett *et al.*¹. The biological age and sexuality shape continua could be compared to determine the correlation between the two signals.

The biological determinant of the signal in such experiments would be unambiguous, isolating the psychological component of subjects' reactions to faces examined by Perrett *et al.*¹, and supporting an argument that is less dependent on questionable semantics.

David A. Meyer*, Melanie W Quong†

*Project in Geometry and Physics,

Department of Mathematics 0112, and

†Department of Biology 0366,
University of California at San Diego,

La Jolla, California 92093, USA

e-mail: dmeyer@chonji.ucsd.edu

1. Perrett, D. I. *et al.* *Nature* **394**, 884–887 (1998).
2. Brennan, S. E. *Leonardo* **18**, 170–178 (1985).
3. Enquist, M. & Ghirlanda, S. *Nature* **394**, 826–827 (1998).
4. Folstad, I. & Karter, A. J. *Am. Nat.* **139**, 603–622 (1992).
5. Wedekind, C. *Proc. R. Soc. Lond. B* **247**, 169–174 (1992).
6. Hamilton, W. D. & Zuk, M. *Science* **218**, 384–387 (1982).
7. Grossman, C. J. *Science* **227**, 257–261 (1985).
8. Alexander, J. & Stimson, W. H. *Parasitol. Today* **4**, 189–193 (1988).
9. Zahavi, A. J. *Theor. Biol.* **53**, 205–214 (1975).
10. Kirkpatrick, M. & Ryan, M. J. *Nature* **350**, 33–38 (1998).
11. Thornhill, R. & Gangestad, S. W. *Hum. Nature* **4**, 237–269 (1993).

Perrett *et al.* reply — Our computer graphic manipulations of the 'geometrical' differences between male and female face shapes generate stimuli that embody the 'psychological' meaning of masculinity and femininity¹. Meyer and Quong question whether our stimuli accurately imitate the 'biological' effect of sex hormones on facial structure, arguing that this should be defined from same-sex individuals whose levels of sex hormones differ.

As visible sexual dimorphism in the human face increases rapidly at puberty when sex steroids influence tissue growth, our working assumption is that the geometrical differences between male and female faces parallel the differences between individuals with high and low androgen levels. Masculinity in mammals is testosterone dependent, whereas femininity represents a more neutral developmental state². Individuals with an XY chromosome pattern and complete androgen insensitivity syndrome have a female appearance. It is unlikely that geometrically 'feminizing' a face produces facial attributes associated with higher than average levels of androgens.

Meyer and Quong suggest that our subjects may be attracted to cues of youth rather than femininity in faces. Cues to femininity and youth are inextricably related, however, because of sex differences in the amount of facial growth at puberty.

Pubescent male faces grow a great deal, whereas female faces change less and therefore remain relatively child-like. As we have shown¹, manipulations based on geometric sex differences in face shape alone affect apparent age. Certainly, face stimuli derived from males with measured high and low levels of testosterone would reflect androgenic effects on target tissues more accurately than our stimuli. Nevertheless, high-testosterone composite faces will carry more 'mature' features and may still appear older than low-testosterone composites.

Evolutionary theories do not predict that females should favour youthful-looking males; instead, they predict that cues to high male status, such as dominance, will be considered attractive³. Our masculinized male faces were rated as more dominant than feminized or average facial stimuli. Facial dominance is a predictor of male status in at least some human hierarchies⁴ and yet subjects in our experiments preferred subdominant male faces. It is unclear whether this is due to a preference for youthfulness or a preference for other characteristics attributed to femininity (such as apparent warmth, kindness or a willingness to invest in progeny). What is apparent is that explanations of male attractiveness based on female sensory biases⁵ or honest Zahavian handicaps fail to predict the preference for slightly feminized male facial characteristics.

Our investigation indicates that females select mates on characteristics other than, or in addition to, dominance.

D. I. Perrett, I. Penton-Voak

School of Psychology, University of St Andrews,

Fife KY16 9JU, UK

e-mail: dp@st-and.ac.uk

1. Perrett, D. I. *et al.* *Nature* **394**, 884–887 (1998).
2. Owens, I. P. F. & Short, R. V. *Trends Evol. Ecol.* **10**, 44–47 (1995).
3. Buss, D. M. *Behav. Brain Sci.* **12**, 1–49 (1989).
4. Mueller, U. & Mazur, A. *Behav. Ecol.* **8**, 569–579 (1997).
5. Enquist, M. & Ghirlanda, S. *Nature* **394**, 826–827 (1998).

errata

In Christopher Austin's article "Lizards took express train to Polynesia" (*Nature* **397**, 113–114; 1999), the circles in Fig. 1 should have been blue and the squares should have been red, to be consistent with the figure legend.

The article "Josephson effect and a π -state in superfluid ^3He " by Olivier Avenel, Yuri Mukharsky and Eric Varoquaux (*Nature* **397**, 484–485; 1999) should have contained the following reference list:

1. Pereverzev, S. V., Loshak, A., Backhaus, S., Davis, J. C. & Packard, R. E. *Nature* **388**, 449–451 (1997).
2. Backhaus, S., Pereverzev, S., Loshak, A., Davis, J. C. & Packard, R. E. *Science* **278**, 1435–1438 (1997).
3. Backhaus, S. *et al.* *Nature* **392**, 687–690 (1998).
4. Avenel, O. & Varoquaux, E. *Jpn. J. Appl. Phys.* **26**, (suppl.) 1798–1803 (1987).
5. Avenel, O. & Varoquaux, E. *Phys. Rev. Lett.* **60**, 416–419 (1988).

Exit *Homo sapiens*, stage left

The Age of Spiritual Machines:
When Computers Exceed Human
Intelligence

by Ray Kurzweil

Viking: 1998. 376 pp. \$25.95, £15.69

Robot: Mere Machine to
Transcendent Mind

by Hans Moravec

Oxford University Press: 1998. 220 pp. \$25,
£15.12

John L. Casti

According to evolutionary biologists and palaeontologists, about 99 per cent of all species that have ever inhabited the Earth are now extinct. If you believe the thesis of these two books, *Homo sapiens'* day in the sun as the leading intellectual force on this planet is also just about over. In fact, Kurzweil and Moravec both turn out the light for humans somewhere in the middle of the twenty-first century, at what Kurzweil would call the "dawning of the age of spiritual machines".

Kurzweil, one of the great innovators in computing technology, predicts that computers exceeding the memory capacity and computational power of the human brain will be available within the next 20 years. In the scenario sketched out in *The Age of Spiritual Machines*, well before the year 2050, information will be fed directly into our brains via neural connections with machines, and the distinction between humans and machines will be totally blurred. At that point, computers will have acquired consciousness, human-style, and will embark on an evolutionary pathway that diverges from our own in almost every way that counts — physically, emotionally and cognitively.

In the world of artificial intelligence, such behaviour would be termed 'strong AI'. Kurzweil's vision is the strongest of strong AI — with a vengeance. Lest you fear that machines will then just keep humans around for laughs and/or as pets, Kurzweil predicts that humans will gradually co-evolve with these machines via neural implants that enable us to 'upload' our carbon-based neural circuitry into whatever kind of hardware machines are using in the late twenty-first century. At this point, there will be no clear distinction between humans and computers, and life expectancy will no longer be a term that pertains to intelligent beings. So humans won't exactly disappear; they will simply merge with the machines. Isn't that comforting?

The rationale underlying the arguments of both books is what Kurzweil has termed "The law of accelerating returns", which applies specifically to evolutionary processes like the development of computer technology. Basically, this states that as order exponen-



Man made machine: super-intelligent robots 'will be our heirs, sharing our goals and values'.

tially increases, time exponentially speeds up. In other words, the time interval between notable events gets shorter as time passes. The crucial implication is then that the valuable products of the evolutionary process also accelerate. Since computer technology is an evolutionary process that builds on its own progress, the time required to accomplish a fixed objective gets exponentially shorter as time goes on. So, for instance, it took 90 years for the first million instructions per second (MIP) to be achieved versus one day for an additional MIP now. If you buy into this law — and all empirical evidence currently available supports it completely — then the replacement of humans by machines as the primary intellectual force on Earth is indeed imminent.

A crucial part of the argument is that the computational process is evolutionary, so it will not crash into the same kind of barriers faced by other sorts of exponentially growing processes, such as Moore's law (which states that the surface area of an integrated circuit chip is halved every 12 months).

How will computing power continue to accelerate as Moore's law dies out? According to Kurzweil, it will use the third dimension of space in semiconductor design — including circuits that don't generate heat. And as this technology tops out, others like DNA and quantum computing will step in to take its place. Thus, the entire evolutionary process of computing is literally unbounded. So goes Kurzweil's argument, anyway.

The Moravec volume draws essentially the same conclusions, emphasizing the emergence of robots embodying the greatly enhanced computing technology of the twenty-first century. But his vision of the "robotic takeover" is not an apocalyptic one; rather, he takes the startling view that "Intelligent machines, which will grow from us, learn our skills, and share our goals and values, can be viewed as children of our minds". In other words, these super-intelligent robots will be our heirs, and as such we will want them to outdistance us, much as parents want their biological heirs to have better, more productive lives than they have themselves.

Moravec then goes on to argue that these intelligent machines offer us lowly carbon-based forms the best chance we'll ever get for immortality by uploading ourselves into advanced robots — the very same process that Kurzweil emphasizes. At this point, we will become our children and live for ever. This is either a science-fiction nightmare or a utopian fantasy: take your pick. In either case, the story Moravec weaves is fascinating.

While each of these books tells much the same tale, they do it in very different ways. Kurzweil's account focuses mostly on the intelligence of machines and the way humans will interface with them via direct neural connections to our brains. The writing is always lively and the arguments are well presented and amplified by numerous boxed cut-outs, tables and graphs showing the ever-accelerating pace of computing technology. The Moravec volume, on the other hand, spends a considerable number of pages on a detailed discussion of the evolution of robots, before moving on to some of the same types of speculations about machine intelligence as Kurzweil. The writing is equally lively, bolstered with diagrams and charts to support the arguments.

I recommend both books to any reader who wants a mind-expanding account of the rise of the age of intelligent machines. These

two volumes are nothing less than a blueprint for how to shove *Homo sapiens* off centre-stage in evolution's endless play. □

John L. Casti is at the Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501, USA.

Affecting climate in a second-best world

Human Choice and Climate Change, vols 1–4

edited by Steve Rayner and Elizabeth L. Malone

Battelle: 1998. 1,714 pp. \$250, £150 (hbk); \$100, £60 (pbk)

Paul Ekins

If the importance of climate change is related to the quantity of words that have been written about it, these volumes would suggest it is easily the most significant of contemporary environmental issues. *Human Choice and Climate Change* runs to four volumes and 1,400 pages of text, with chapters that stretch up to 80 pages and have well over 100 references each.

This ambitious project aims to present an alternative view of climate change to that provided by the Intergovernmental Panel on Climate Change (IPCC) — which, of course, has also been responsible for some weighty tomes. The questions underlying the IPCC approach may be characterized as: Is anthropogenic climate change occurring and how might it develop? What might be the impacts on ecosystems and human societies? How should human societies, and the international community, respond?

Human Choice and Climate Change, in contrast, starts with such questions as: How do scientists choose to study climate change and form a scientific consensus? How do people attribute blame for climate change and choose solutions? How are climate-change policy instruments chosen? Why and how did the international community choose to address climate change?

These questions make it clear that *Human Choice and Climate Change* treats the issue of climate change more as a social construction than as a natural-science phenomenon. Its authors occupy a spectrum of acceptance of what the IPCC natural scientists say about climate change. At the most sceptical end, one chapter by Michael Thompson and Steve Rayner refers to climate change as part of the “hegemonic myth” of “global vulnerability and fragility”. It suggests that a “potentially implausible hypothesis that human activities will warm the entire planet” has been picked up by the public because of “the lay propensity to attribute changes in weather patterns to human activity” rather than because of any scientific evidence.

At the other end of the spectrum is a chapter by Donald J. Wuebbles and Norman J. Rosenberg presenting the natural science of climate change very much as the IPCC does. Similarly positioned, another chapter assesses the possible impacts of climate change on land and water use, coastal zones and oceans, and on people, and discusses possible response strategies. Another explores how the world's energy systems have developed to their current state, how they might evolve, and the costs involved, as a result both of climate change and of attempts to mitigate it. Another discusses the possibilities, challenges and limitations of integrated assessment modelling. None of these chapters would be out of place in an IPCC report; they all make a useful contribution to knowledge about the issues.

In the last volume the editors try to pull together all the insights and present them in a form that would be useful to people who make policy. The final chapter is even called ‘Ten suggestions for policy-makers’, the project's equivalent to the IPCC's policy-makers' summaries.

One of the main points in this volume is the distinction drawn between two styles of social science: the descriptive paradigm that “analyzes social systems in terms of natural science metaphors” and the interpretive approach that “refers to the analysis of the values, meaning and motivation of human agents”. *Human Choice and Climate Change* contains good examples of both.

Rayner and Malone would like to see more research that seeks to integrate these two approaches, and would clearly recommend this as the basis for the climate-change research programme. Yet, on the evidence of these books, integration of the approaches is too strong an aspiration. The tensions and contradictions between them would make any such union chaotic. The descriptive approach depends on an acceptance of an objective reality that it is trying to describe. The interpretive approach sees the world much more in terms of social construction.

Both may generate insights into the human condition. But each approach calls into question the other's fundamental axioms. The integration of such contrasting styles would be unlikely to be useful even if it were possible.

As for policy-makers, my hunch is that those few of them with the time to read these volumes will be more impressed by the descriptive than the interpretive chapters. The problem with the latter from a policy point of view is that they increase substantially the already formidable uncertainties connected with climate change.

The editors' ten recommendations do not advise policy-makers on how much to reduce emissions, or what policy instruments to use. Rather, they urge them to “view the issue holistically”, “recognise the limits of rational planning” and “employ the full range of analytic approaches and decision aids”. This might be good advice in a policy utopia, but the danger is that, in a second-best world, taking it too literally could bring policy-making to a complete halt. And if the IPCC has got its descriptive science right, that could be a most unfortunate result for the good intentions and high scholarship that have gone into this work. □

Paul Ekins is in the Department of Environmental Social Sciences, Keele University, Staffordshire ST5 5BG, UK.

Happy accidents?

The Pattern of Evolution
by Niles Eldredge

W. H. Freeman: 1999. 250 pp. \$24.95, £17.95

Mark Pagel

The palaeobiologists who peered into the fossil records in the early part of this century were not prepared for what they saw. Pre-Darwinian palaeobiology had been dominated by the need for explanations consonant with the Creator's plan. The first hint of change came from the geologist Charles

Diversity in danger

Relish the sight of this anglehead lizard while you can. It is one of the 163 species of reptile and more than 510 species of land animals that exist only on the Philippine islands, and which are in grave danger from the

economic pressures on the Filipino people to plunder the precious forest. *Vanishing Treasures of the Philippine Rain Forest* by Lawrence R. Heaney and Jacinto C. Regalado Jr (University of Chicago Press, \$24, £19.25, pbk) describes the richness and vulnerability of this “hottest of all hotspots”.



Lyell, who dismayed sensible people of the 1830s by proclaiming that the natural world was neither static — as God left it — nor young. Later, Darwin proclaimed a wholly material explanation for species, based on the principle of descent with modification. Lyell had opened the door, and Darwin showed God out.

Palaeobiologists flocked to these scientific visions of a world in a constant state of flux and admixture. But instead of finding the slow, smooth and progressive changes Lyell and Darwin had expected, they saw in the fossil records rapid bursts of change, new species appearing seemingly out of nowhere and then remaining unchanged for millions of years — patterns hauntingly reminiscent of creation.

But there was no turning back, and biologists have for the past century fought over how best to explain the diversity of life. Does it arise from the gentle moulding of genes by natural selection, or are there powerful, invisible hands at work that visit revolutionary changes upon the biotic world? Niles Eldredge has for three decades been among the leading voices of dissent against what he regards as a simplistic “ultra-Darwinian” view of the world: that the struggle among genes for reproductive success provides a sufficient explanation for organismic diversity.

The Pattern of Evolution is his passionate account of how he came to that view. Eldredge outlines in a popular style a theory of ‘matter-in-motion’ to entrain biological evolution to the many climatic, geological and even tectonic forces that sculpt the Earth. Eldredge sees their power, inevitability and even occasional cyclicity behind many of the otherwise puzzling patterns that characterize the history of biotic evolution on this planet.

A favourite textbook example is the dinosaurs. Not until they became extinct did the primitive mammals of the time radiate into the diverse forms we know today, including ourselves. Dinosaur extinction may have been an accident — a meteor struck the Earth; *ergo*, it is argued, we are happy accidents as well. Does this ‘wonderful life’, to paraphrase Stephen Jay Gould, one of Eldredge’s collaborators over the years, derive more than is appreciated from contingency and luck?

Late in the twentieth century we should be asking if a synthesis will ever emerge from this old and decidedly Hegelian conflict of thesis and anti-thesis. Before it does, a ‘third way’ (but not a ‘*neue mitte*’) may demand an accounting. Whole genes can emerge *de novo* from the process of gene duplication. The major groups of animals — such as invertebrates versus vertebrates, or the fish versus the mammals — have different numbers of the important and fundamental Hox genes that determine their very divergent pheno-



George Lyell: contributed to the revelation that the Earth is in a constant state of flux.

types. It is as if the extra genes give rise to the way of life. No one knows how general this is. Perhaps another powerful, stochastic and relentless driving force of evolution resides right inside our genomes. □

Mark Pagel is at the School of Animal and Microbial Sciences, University of Reading, Whiteknights, Reading RG5 6AH, UK.

Species without frontiers

Life out of Bounds: Bioinvasion in a Borderless World

by Chris Bright

W. W. Norton/Earthscan: 1998. 287 pp.

\$13/£12.95 (pbk)

Harold A. Mooney

The threat of biological warfare has resulted in extraordinary measures by nations to detect and defend themselves against potential weapons of destruction. But inadvertent biological invasions have actually been occurring for centuries, decimating human populations and the crops and livestock upon which they depend. The fundamental driver of this has been the transport of people and their goods, along with biological stowaways, some of which have been benign, others extraordinarily destructive.

For some reason, insufficient attention has been paid to the scale of this interchange, its increasing tempo, and its enormous economic and ecological impact. International commerce has broken down the ancient bio-

geographic barriers that allowed the evolution of the rich biological diversity on which we based our civilizations. The breakdown of these barriers has resulted in a biological scramble of species competing for newly opened resources. This is now happening on a global scale, with serious consequences for the future course of evolution and the number of species that will populate the Earth.

This is pretty dramatic stuff — and yet the forces to counter this tidal wave have not yet been assembled. The first part of the job of resistance is to raise awareness of the problem and of the potential courses of action available to deal with it. Chris Bright has done an admirable job of marshalling the facts and writing about them in a clear and compelling manner.

This book developed from an excellent article he wrote in 1995 for *World Watch*, on the spread of exotic species. Since then, a landmark conference on alien species held in Norway in 1996 has provided new information as well as the beginnings of an international plan to address this complex and urgent issue. Much of this information appears in this well-documented book.

The book is in three sections. First, Bright surveys the geography of invasions and shows that none of the major biotic realms — fields, forests, waters and the special case of islands — is now free of the impact of invaders from other continents. He then describes how the change described by some as the new geological epoch of biological homogeneity, the Homocene, came about. Both intentionally and by accident, human activity has mixed up the biota of the Earth, incurring large economic and environmental costs in the process. Finally, a course of action to deal with this crisis is presented.

In the section on the geography of invasions, Bright provides case after well-documented case of invading species that have had a profound impact on human societies. He describes, for example, the dramatic story of Lake Victoria, which may now be in a state of ‘chronic emergency’ as a result of successive waves of invasive species. The first of these was the Nile perch, introduced in 1962, which drove a couple of hundred native fish species to extinction and indirectly altered the surrounding forested ecosystems as local firewood was used by fishermen to dry the very large Nile perch. These major alterations apparently led to the successful establishment of water hyacinth in 1990. This invader is in turn making the shoreline anoxic, clogging pumping stations and severely affecting fishing activity. It is also apparently increasing the incidence of snail- and mosquito-borne diseases. This chain of events has affected millions of people.

The ‘villain’ species in such cases are either purposeful or accidental introduc-



Plant-locked: the spread of the water hyacinth in Lake Victoria has closed small ports and harbours.

tions. Cases of accidental introductions are growing because of increased shipping and the use of ballast water. Ballast water, acquired in one port and dumped in another, brings large numbers of live animals to new habitats. Around 1982, a small comb jellyfish was brought in this way from the east coast of the Americas to the Black Sea. This organism became well established and devastated the fisheries, affecting two million people. In a poor exchange, the zebra mussel was apparently taken in ballast water from the Caspian Sea to the Great Lakes of North America around 1986. This mussel is doing so well in its new habitat and its numbers are now so high that it controls the ecological dynamics of many areas, and adversely affects waterworks such that remediation is costing billions of dollars.

Doing something about invasive species will not be easy, as Bright points out, because we generally do not know which organisms will become successful invaders, where and when invasions will occur, or what effects they will have. If ever there was a rationale for the precautionary principle on the release of any biotic material into a new ecosystem, this is it. The policy response to the threat of invasions has been weak and uncoordinated, and only the most devastating invaders get serious attention. Bright thinks we need to strengthen the treaties relevant to invasive species, develop codes of conduct for industries involved in the trade of biotic material, create an information clearing-house, improve our capabilities of biological control, and increase ecological literacy. All of these are important and achievable goals.

There is some hope of providing an international focus on the problem, and for the development of tools and determination

to deal with this issue. The Scientific Committee on Problems of the Environment, and its partners the World Conservation Union (IUCN), the United Nations Environment Programme and the Commonwealth Agricultural Bureau International, have launched a Global Invasive Species Programme to provide the means for recognizing and dealing with this pervasive problem. Of particular concern is the threat it poses to species diversity and ecosystem services and to human welfare in general.

Bright's book is a clarion call to action. It is exceptionally good value in every sense. Every biologist should be well versed on this issue, and reading this book would be an excellent starting point. □

Harold A. Mooney is in the Department of Biological Sciences, Stanford University, Stanford, California 94305, USA.

Ritual meanings lost in a vegetable stew

Plants of Life, Plants of Death

by Frederick J. Simoons

University of Wisconsin Press: 1999. 568 pp.

\$65, £57.95 (hbk), \$34.95, £31.50 (pbk)

Roy Vickery

Mankind's attitude to the ritual use of plants arose primarily from basic human concerns about life, fertility and death, rather than medicinal or nutritional factors — according to Frederick Simoons, at least.

Simoons, of the University of California, seeks to demonstrate this by accumulating a vast amount of information on holy basil (*Ocimum sanctum*), sacred figs (*Ficus religiosa* and *F. indica*), mandrake (*Mandragora*

officinarum), ginseng (*Panax ginseng*), garlic (*Allium sativum*), urd bean (*Vigna mungo*) and broad (or fava) bean (*Vicia faba*).

His work is library-based in the tradition of James Frazer's *The Golden Bough*; little, if any fieldwork appears to have been undertaken. Of the book's 568 pages, 230 are devoted to notes and bibliography. On the positive side he writes in a clear, jargon-free style. On the negative side, like Frazer, he accumulates so much information that the reader is overwhelmed. In the earlier chapters Simoons himself seems overwhelmed, to the extent that he is reluctant to draw conclusions from his assembled evidence. Perhaps conclusions are there, but they are hard to find.

Simoons is right to challenge the currently prevalent view that mankind's relationship with plants is derived from a plant's economic, nutritive or medicinal value, or its ability to harm. For example, some people did develop favism after consuming broad beans, but this alone was not responsible for the Pythagorean ban on beans: there were more complex cultural reasons. However, his case could have been presented in a more succinct and convincing manner.

A more disciplined approach, which limited the amount of information and gave greater prominence to his hypothesis, would have improved the work and made it more useful. As it is, only the most determined and alert readers will successfully negotiate their way through the often repetitive evidence to reach the valid, but frequently somewhat unrelated, findings. □

Roy Vickery is in the Department of Botany, The Natural History Museum, London SW7 5BD, UK.



Cultivated by culture: garlic (top) and mandrake became popular for their ritualistic use.

Signals of need in parent–offspring communication and their exploitation by the common cuckoo

R. M. Kilner, D. G. Noble & N. B. Davies

Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK

Nestling birds present vivid gapes and produce loud calls as they solicit food, but the complexity of the display is poorly understood. Here we explain the function of reed warbler begging signals and show how they are exploited by the common cuckoo, *Cuculus canorus*, a brood parasite. Reed warbler parents integrate visual and vocal signals from their young to adjust their provisioning rates, and the two signals convey more accurate information about offspring need than either does alone. The cuckoo chick has a particularly striking begging display which has been suggested to be irresistible to host parents. However, we show that the cuckoo, reared alone in the nest, presents a deficient visual display, and elicits the same amount of care as a reed warbler brood only by compensating with its exaggerated vocal display. Therefore the cuckoo succeeds not through mimicry of the host brood begging signals, but by tuning into the sensory predispositions of its hosts.

Many animal displays are highly complex^{1,2}, often built from several component signals of different sensory types^{3,4}. For example, nestling birds beg with a combination of brightly coloured gapes⁵ and repetitive begging calls⁶. Despite their ubiquity, and a range of theoretical models^{7–11}, empirical understanding of multicomponent displays is limited¹². Here we explain why the nestling begging displays of reed warblers (*Acrocephalus scirpaceus*) and their brood parasite, the common cuckoo (*Cuculus canorus*), include both visual and vocal components. Our experiments show that reed warbler parents use both signals to adjust their provisioning rates at the nest. Together, these signals provide parents with more information about their offsprings' need than is available from either signal alone. Moreover, we show that reed warblers follow exactly the same integration rule when provisioning a single cuckoo in their nest as when feeding a brood of their own young. The way in which parents integrate two nestling signals illuminates the function of the cuckoo's extraordinarily rapid begging call¹³. Rather than whipping parents into a feeding frenzy at the nest^{14,15}, the supernormal begging call instead serves to persuade parents to work at their normal rate by compensating exactly for the cuckoo's subnormal visual display of a single gape.

Information carried by begging signals

We studied a population of reed warblers at Wicken Fen, Cambridgeshire, UK, in the summers of 1996–1998. Nestling reed warblers solicit food from their parents by exhibiting a bright yellow gape and calling repeatedly. Although previous work with other species has established that the various components of the begging display are individually correlated with nutritional state^{5,16,17}, it is not known how they combine to signal food requirements. We investigated whether gape area and begging call rate convey distinct information about the extent of food deprivation (see Methods). Chicks were temporarily removed from the nest at a range of different ages (day 3/4 and day 6/7) in broods of four (the modal brood size). They were fed with chick-rearing mix until they stopped begging, and then kept warm but unfed. Every 10 min, for the next 110 min, we stimulated the brood to beg, recording the begging behaviours on video and audiotape (see Methods). We found that total gape area displayed and begging call rate each varied significantly with both the extent of food deprivation and chick age (Fig. 1). Hence, the different elements of the begging display carried 'multiple messages'¹¹, each element reflecting both age and

hunger^{10,11}. Moreover, because the two elements in the begging display each explained different components of variance, for both chick age and food deprivation, they also functioned as 'back-up' signals (Fig. 1)¹¹.

We checked the robustness of our conclusions by using a separate analysis of a larger data set with a different measure of offspring need, namely the amount of food required to cause the nestling reed warblers to stop begging. Nestlings were temporarily removed from the nest at a range of different ages (day 2 to day 8) and in broods of one ($n = 37$) or four ($n = 12$). They were fed with chick-rearing mix until they stopped begging, and again kept warm but unfed. After 110 min, we induced the brood to beg and recorded the begging behaviours on video and audiotape. Finally, we scored the amount of food eaten by the brood until begging stopped (see Methods). To explain the variance in the amount of food consumed, we used a multiple regression analysis, with total gape area displayed and begging call rate as the two independent factors. We found that each signal explained different components of variance in the broods' nutritional condition. There were significant independent effects of gape area ($t_{48} = 5.63$, $P < 0.0001$) and call rate ($t_{48} = 2.02$, $P = 0.042$) on the amount of food consumed. The regression equation derived was: food consumed (in ml) = 0.002 (gape area displayed (in mm²)) + 0.009 (number of calls per 6 s) – 0.011 , and this explained 77.6% of the variance in food consumption. We conclude that multiple begging signals provide parents with a more accurate assessment of offspring need than can be determined from either signal alone.

Parental response to multiple signals

Do parents regulate their provisioning rate in relation to both the rate at which the brood calls and the total gape area displayed? We tested this possibility by subjecting the parents of broods of two and four nestlings, aged 6/7 days old, to playback of begging calls which were broadcast through a small speaker, attached to the side of the nest, every time the nestlings gaped for food. The three treatments were: no playback, playback of one chick calling and playback of four nestlings calling (see Methods). Clearly, the effects of these playbacks on the parents would be confounded if they affected the behaviour of the nestlings. To examine this, we recorded the duration of the brood begging display during each treatment, because in another species this is influenced by sibling vocalizations¹⁸. There was no indication that the playback of

begging calls affected this aspect of chick begging behaviour ($F_{2,32} = 1.17$, $P = 0.90$), nor did playback alter the likelihood of gaping (Friedman analysis of variance (ANOVA); $\chi^2_2 = 1.63$, $n = 12$, $P = 0.43$). Thus by broadcasting begging calls as nestlings displayed, we augmented only the broods' vocal displays (see also below), whereas the manipulations of brood size affected both visual and vocal signals equally. There was a significant effect of both brood size and playback treatment on the rate at which parents delivered food to the nest, and the parental response to playback was independent of brood size (Fig. 2). We have established previously that provisioning rate correlates with nestling weight gain¹³.

In a different analysis, we used data from 59 brood-size manipulations and 52 playback manipulations (see Methods) to investigate the separate effects of visual and vocal signals on provisioning rates to chicks aged 6/7 days old. We used field recordings of broods manipulated to be different sizes to check that the calling rate of individual chicks was not affected by altering the brood size. A strong positive linear relationship between brood size and brood call rate ($R^2 = 0.90$, $F_{1,15} = 151.6$, $P < 0.0001$), with a small inter-

cept (1.09), indicated that the mean begging call rate per chick did not vary with brood size. To explain the variance in the rate of parental provisioning, we used a multiple regression analysis with the maximum number of chicks calling per nest (the number of chicks in the brood plus the number of chicks calling on the playback tape) and the maximum number of gapes on display per nest (the number of chicks in the brood) as the two independent factors. We assumed that the whole brood gaped at every visit. Our limited data for broods of two to four chicks show that the mean proportion of the brood that gaped per visit was 0.8 ($n = 13$), but we have no data for larger brood sizes. As before, we checked that the data fitted the assumptions of a linear multiple regression model.

We found that reed warbler parents combined multiple signals from their young to adjust their provisioning rate at the nest. There were significant independent effects of the maximum number of gapes displayed per nest ($t_{110} = 4.42$, $P < 0.0001$) and the maximum number of chicks calling per nest ($t_{110} = 5.69$, $P < 0.0001$) on the number of feeds delivered by both parents per hour. Together, both signals explained 62.7% of the variance in nest visit rate ($F_{2,110} = 93.28$, $P < 0.0001$). The regression equation derived was: feeds delivered per hour = 2.28 (maximum number of gapes displayed) + 2.30 (maximum number of chicks calling) + 8.23. A polynomial regression does not provide a better fit than the linear multiple regression, which indicates that parents probably did not reach their provisioning limit within the range of our experimental treatments.

The regression equation explains parental provisioning rate in terms of the number of chicks displaying visually and vocally 6/7 days after hatching. We studied whether the equation could be expressed more generally to explain provisioning rates to reed warbler nestlings of different ages and in different brood sizes. To generalize the equation from day-6/7 chicks to young of all ages, we

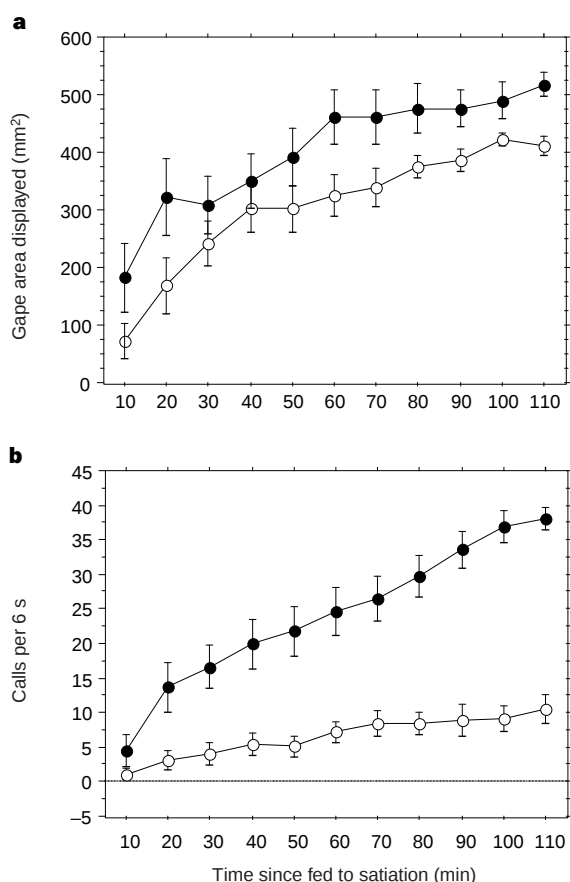


Figure 1 Effect of food deprivation. Effect of food deprivation on **a**, the total gape area displayed by all chicks (see Methods) and **b**, the rate of calling by broods of four reed warblers, on day 3/4 (open circles; $n = 9$) and day 6/7 (filled circles; $n = 10$). Plots show means $\pm$ s.e.m. The data were analysed with a repeated-measures ANOVA, with two within-factors (food deprivation time and type of begging signal) and one between-factor (chick age). There was a significant effect of food deprivation ($F_{10,170} = 28.94$, $P < 0.0001$) on both begging displays. Moreover, the rate of change in gape area displayed with increasing food deprivation time was significantly greater than the rate of change in begging call rate ($F_{10,170} = 24.80$, $P < 0.0001$). Older chicks displayed more gape area and called at a higher rate than younger chicks ($F_{1,17} = 6.47$, $P = 0.021$), and begging call rate changed more with increasing age than did gape area displayed ($F_{1,17} = 3.88$, $P = 0.065$, power = 0.448).

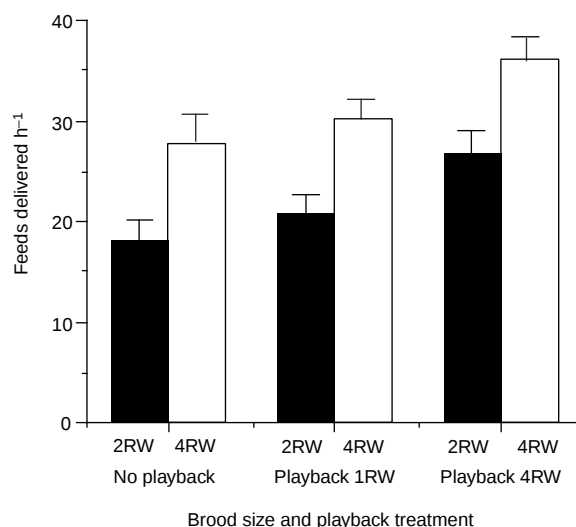


Figure 2 Effect of playback treatments. The effects of the playback treatments at broods of two (2RW; $n = 10$ pairs) and four (4RW; $n = 11$ pairs) chicks on feeds delivered per hour. Bars represent means $\pm$ s.e.m. Each pair of parents was subjected to three treatments: no playback, playback of one reed warbler chick calling (playback 1RW) and playback of four reed warbler chicks calling (playback 4RW). The data were analysed with a repeated-measures ANOVA with one within-factor (playback treatment) and one between-factor (brood size). There was a significant effect of brood size ($F_{1,19} = 13.85$, $P = 0.0014$) and playback treatment ($F_{2,38} = 13.42$, $P < 0.0001$) on feeds delivered by both parents per hour, but no interaction between the two ($F_{2,38} = 0.041$, $P = 0.96$). In a separate analysis, there was no significant effect of pair identity on the response to the treatments with either broods of two chicks ($F_{9,29} = 0.88$, $P = 0.56$) or broods of four chicks ($F_{10,32} = 1.90$, $P = 0.10$).

divided the 'gape' coefficient by the mean gape area of a day-6/7 reed warbler chick (140 mm^2), and divided the 'calls' coefficient by the mean number of calls per 6 s given by an unmanipulated day-6/7 reed warbler chick in the field (12.95 calls/6 s). The regression equation thus derived was: feeds delivered per hour = $0.0162 (\text{gape area displayed (in } \text{mm}^2)) + 0.178 (\text{calls per 6 s}) + 8.23$. To assess the accuracy of this more general equation, in the field we measured the begging calls produced and gape area displayed by reed warbler young of different ages (day 1 to day 8), in broods of two, three or four, and calculated mean values for broods of each size and age (see Methods). We substituted these means into the regression equation to calculate nine predicted provisioning rates in total. Using data that were not included in the original multiple regression analysis, we compared these predicted values with mean field observations of provisioning rates to broods of two ($n = 1$ day 5 young), three ($n = 2$ day 1/2; $n = 4$ day 3/4; $n = 9$ day 5/6; $n = 4$ day 7/8) and four ($n = 1$ day 1/2; $n = 7$ day 3/4; $n = 10$ day 5/6; $n = 9$ day 7/8) young. If the observations were to match our predictions exactly, the regression of the former on the latter should not differ significantly from the line $y = x$. This is exactly what we found ($t_7 = 1.43$, $P > 0.1$).

It may seem remarkable that parental provisioning rates for broods of one to eight young, from one to eight days after hatching, can be explained simply as a response to the gape area displayed and the brood calling rate. However, the results of our experiments with chicks in the laboratory indicate that both signals can convey accurate information about chick age, brood size and nutritional state, so perhaps it is not surprising that two signals alone can explain so much of the variance in parental provisioning rate. We suggest that the visual display may provide parents with a crude estimate of how frequently to feed their young, because it is correlated with brood size and chick age, whereas the vocal display may enable parents to fine-tune their provisioning in relation to the brood's level of hunger, because this signal provides more finely graded information about nutritional need. For example, in day-6/7

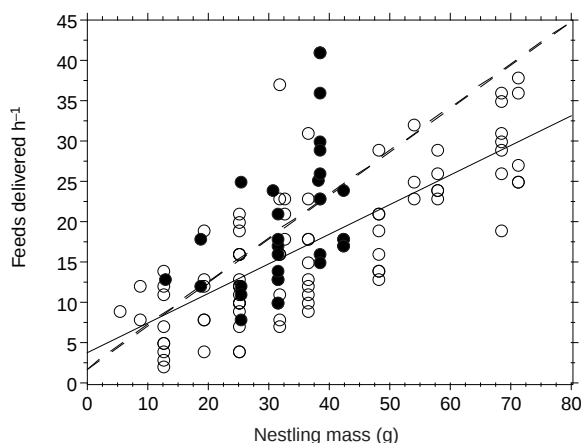


Figure 3 Relationship between nestling mass and feeding rate. The relationship between the mass of young in the nest and the number of feeds delivered per hour by both parents to the nest, for cuckoos (open circles, solid line, $n = 88$ observations of 36 different cuckoo chicks fed by 44 different pairs of hosts; some cuckoo chicks were moved between nests) and broods of four reed warblers (filled circles, dashed line, $n = 34$ different broods fed by their own parents). There was no significant difference between either the elevations ($F_{1,118} = 0.15$, $P = 0.70$) or the slopes ($F_{1,118} = 1.30$, $P = 0.26$) of the regression lines for each species. Repeating the analysis using only one randomly selected data point per cuckoo still yielded no significant difference in either the elevations ($F_{1,66} = 0.45$, $P = 0.51$) or the slopes ($F_{1,66} = 1.27$, $P = 0.26$). Data were collected in 1985–1986 and 1996–1998 either by using a video camera ($n = 24$ cuckoo observations) or by observing the nest from a distance through binoculars ($n = 64$ cuckoo and 34 reed warbler observations).

chicks the visual signal reaches a plateau 60 min after being fed to satiation, whereas the vocal signal continues to change (Fig. 1).

In theory, the amount of food supplied by parents to young is likely to be the source of a conflict of interests between the two parties, because investment patterns that are optimal for parents can be suboptimal for offspring^{19,20}. Recent theoretical work has suggested that the conflict may be resolved in the parents' favour if provisioning rates are based on costly, and hence reliable, nestling signals of need^{21,22}. The two regression equations that we have derived experimentally (first to relate chick need to their begging signals, and second to relate the begging signals to parental provisioning) allow us to link chick need quantitatively with parental provisioning rates. In principle, this provides a way in which to measure the outcome of parent–offspring conflict, and to test the prediction that parents supply chicks with exactly the food that they demand. At present, this analysis is beyond our reach, because the units of each regression equation are not the same.

Exploitation by cuckoo chicks

The reed warblers at Wicken Fen are parasitized by the common cuckoo, a brood parasite that relies entirely on its hosts to incubate

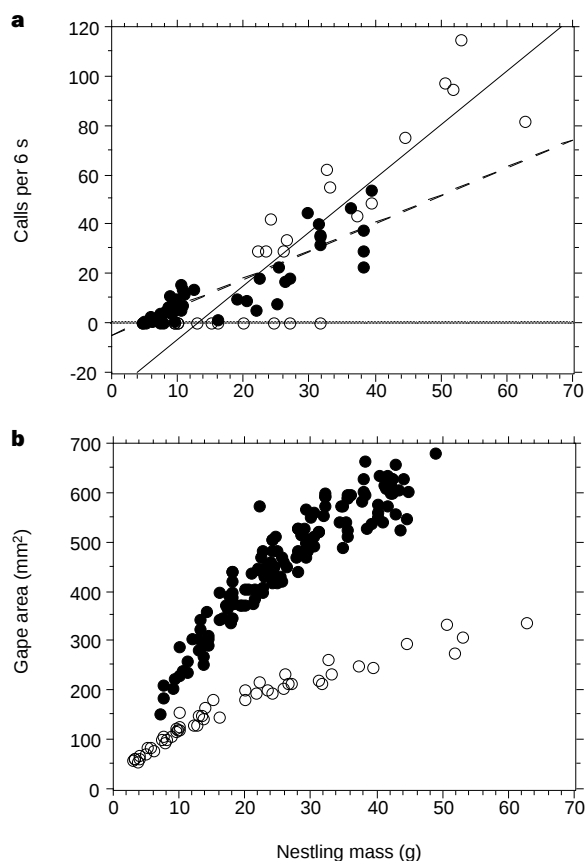


Figure 4 Relationship between nestling mass and begging call rate or gape area. The relationship between nestling mass and **a**, begging call rate, for cuckoos ($n = 25$ measurements) and broods of one and four reed warblers ($n = 49$), or **b**, gape area, for cuckoos ($n = 54$ measurements) and broods of one and four reed warblers ($n = 145$ measurements). Open circles, solid line, cuckoo data; filled circles, dashed line, reed warbler data. Begging displays were measured under standardized conditions of food deprivation in the laboratory (see Methods). The rate of calling by a cuckoo nestling increased at a significantly faster rate with mass than the rate of calling by a brood of four reed warblers ($F_{1,71} = 24.8$, $P < 0.0001$). We compared the rate of increase in gape area with mass in cuckoos and reed warblers by log-transforming the data, and comparing the elevations of the two resulting regression lines. The gape area displayed by a cuckoo nestling grew at a significantly slower rate than the gape area displayed by a brood of reed warbler chicks ($F_{1,188} = 91.7$, $P < 0.0001$).

its eggs and rear its young to independence²³. Shortly after hatching, the cuckoo nestling evicts all of the reed warbler eggs and chicks from the nest²⁴, and so becomes the sole beneficiary of parental care. Apparently oblivious to the destruction of their own reproductive success, the pair of warblers then feeds the imposter, even as it grows to eight times their own body weight. By 2 weeks of age, the cuckoo overflows the tiny nest, and the warblers seem to risk being devoured themselves as they perch on the cuckoo's back in order to bow deep into the enormous gape with food.

Our observations show that the rate at which parents provision a single cuckoo chick closely matches the rate at which they provision a brood of four of their own young, of equivalent mass, the modal brood size in this species (Fig. 3). The food brought to the cuckoo is also the same as for their own young²⁵. Previous experiments have established that the cuckoo's success at eliciting care cannot be attributed simply to its relatively large size. When we placed in a reed warbler nest a single blackbird (*Turdus merula*) or song thrush (*T. philomelos*) chick of the same mean mass as a cuckoo nestling, the *Turdus* nestlings were fed at a significantly lower rate. However, their rate of being fed improved markedly when their own begging display was augmented by cuckoo nestling begging calls broadcast from a small speaker at the side of the nest¹³.

Analysis of a larger data set than presented previously¹³ shows that the *Turdus* nestlings accompanied by cuckoo begging calls obtained significantly more food than cuckoos (unpaired $t_{22} = 2.22$, $P = 0.037$). However, when accompanied by playback of four reed warbler chicks calling, *Turdus* young were fed less frequently than a brood of four reed warblers (unpaired $t_{17} = 1.99$, $P = 0.06$, power = 0.454). Both of these results can be explained if we take into account the total gape area displayed during begging. *Turdus* nestlings have a significantly greater gape area than cuckoos of the same mass, and broods of four reed warblers display a greater total gape area than *Turdus* young ($F_{2,41} = 30.7$, $P < 0.0001$). These results provide further evidence that parents integrate visual and vocal nestling signals to decide how frequently to feed their young, and again indicate that the main visual signal is the gape area displayed, and that the main vocal signal is begging call rate.

At 6–8 days of age, a cuckoo nestling's begging call closely matches the rate of calling of a brood of four reed warblers¹³. However, as the cuckoo grows, its begging call becomes increasingly more rapid, far faster than the rate of calling achieved by a brood of

four reed warblers (Fig. 4a). Conversely, despite its enormous gape, the total gape area displayed by the single cuckoo is much less than that of a brood of four reed warblers, and the difference becomes more pronounced with increasing chick mass (Fig. 4b). This is an inevitable consequence of the fact that adult cuckoos have relatively small bills in relation to mass compared with adult reed warblers²⁶, so a single cuckoo will never match the gape area of four reed warblers of the same total mass (Fig. 5).

If reed warblers follow the same integration rule when feeding a cuckoo as when feeding their own young, we can use our regression equation to test whether the increasingly supernormal begging call of older cuckoo nestlings serves to compensate exactly for their increasingly subnormal visual stimulus, compared with a brood of four reed warblers. Using the equation: feeds delivered per hour = 0.0162 (gape area displayed (in mm^2)) + 0.178 (calls per 6 s) + 8.23 , we calculated predicted begging call rates for cuckoos. We measured the gape areas of cuckoo nestlings at different stages of development ($n = 26$ measurements of 20 different cuckoos), as well as their hourly provisioning rates by both parents ($n = 74$ observations of 28 cuckoos fed by 36 different pairs) and calculated mean values for cuckoos on days 4–11, 13 and 15 (see Methods). We substituted these mean scores into the regression equation to calculate the predicted begging call rate at different stages of development.

When we compared our predictions with measurements of unmanipulated cuckoo begging call rates in the field, we found no significant difference between the two (Fig. 6). Therefore reed warblers do indeed use exactly the same provisioning rules when feeding cuckoos and when feeding their own brood (Fig. 7). The increase in the cuckoo's calling rate with age is exactly what we would expect if it is to offset the disadvantage of displaying a single gape. Furthermore, the close match between our predictions and observations (Figs 6, 7) indicates that any cues other than gape area and call rate must play a minor role in determining host provisioning rate. For example, our experiments show that gape colour has no effect on provisioning rate²⁷. A key point is that the cuckoo does not mimic the begging displays of a brood of reed warblers precisely, but instead tunes into the way in which the host parents integrate visual and vocal cues from their young²⁸, exaggerating the latter to

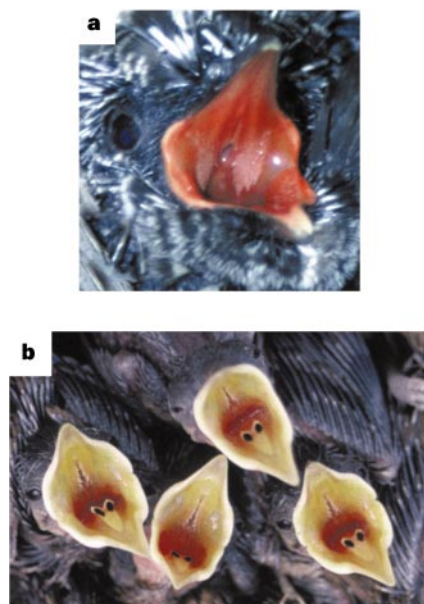


Figure 5 Chick gapes. **a**, A nestling cuckoo, day 11/12; **b**, a brood of four reed warblers, day 6/7. The two pictures are at roughly the same magnification.

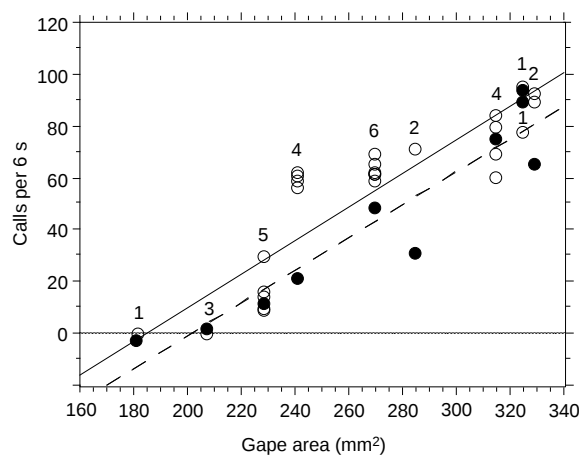


Figure 6 Comparison of observed and predicted begging call rates. Observed cuckoo begging call rate (open circles, solid line) is compared with the predicted cuckoo begging call rate (filled circles, dashed line) that should accompany the display of a particular gape area, assuming that reed warblers use the same rules to feed both a cuckoo nestling and their own young. Numbers refer to sample sizes of cuckoos measured at each age. Both lines increase significantly in relation to gape area ($F_{1,35} = 24.8$, $P < 0.0001$), but there is no significant difference between them in either their slope ($F_{1,35} = 0.018$, $P = 0.89$) or their elevation ($F_{1,35} = 0.77$, $P = 0.78$). Each predicted value corresponds to a particular day of the cuckoo's nestling period (from left to right, day 4–11, day 13 and day 15).

compensate for deficiencies in the former. There is an analogy here with the way in which males can exploit female sensory systems in mate choice²⁹. This sensory exploitation differs from the deception practised by aggressive mimics such as *Photurus* fireflies and *Bolas* spiders³⁰ which, respectively, mimic exactly the visual or chemical cues of their prey.

Theoretically, we would expect the cuckoo to demand a much higher provisioning rate than a brood of reed warblers, because it has no genetic stake in the future breeding success of the parents^{31,32}. However, far from being an irresistible narcotic lure¹⁴, the cuckoo chick apparently struggles to induce parents to feed it at the rate at which they would provision a brood of their own young. Rather than demanding a higher provisioning rate, the cuckoo instead exploits hosts by forcing them to provide care for longer (17 days in the nest plus 16 days after fledging, compared with 11 days in the nest plus 12 days after fledging for reed warbler young)³³. Parents can certainly be made to work much harder in the short term²⁵, so why don't cuckoos solicit a higher rate of provisioning? It has previously been argued that this would not benefit the cuckoo because it might exhaust hosts before the end of the cuckoo's dependent period²⁵. But the data shown here suggest another explanation. By evicting the reed warbler chicks from the nest at the start of its nestling life, the cuckoo gains the benefit of receiving all the food brought to the nest, but pays the cost of being solely responsible for dictating the rate at which it is provisioned. Constrained by its visual display of a single gape, the cuckoo may be unable to solicit a higher rate of feeding, perhaps because there is an upper limit to the rate at which calls can be produced or perceived.

Other species of brood parasite that are reared alongside the hosts' own young may show a different solution to the same problem. These species might suffer the cost of sibling competition for food at each nest visit^{34,35}, but at least they gain the benefit of assistance in soliciting a high visit rate. In short, the constraints of eliciting a high provisioning rate alone, or of losing food in sibling competition, may mean that we never see brood parasites being fed at the high rate we might predict. □

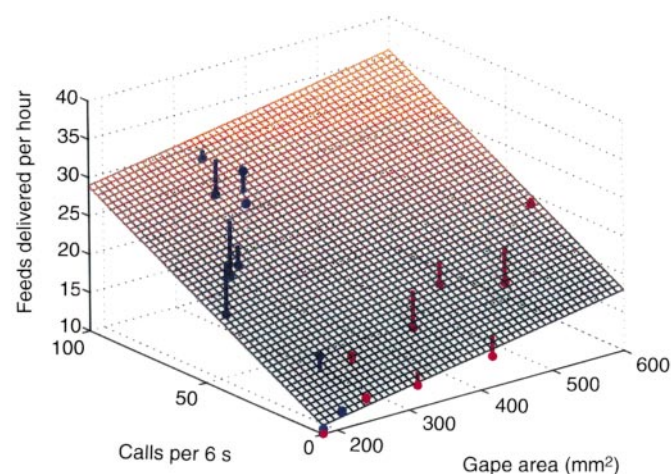


Figure 7 Comparison of predicted and observed provisioning rates. Predicted provisioning rates, derived from our experiments with reed warbler nestlings, are shown by the plane, which represents the regression equation: feeds delivered per hour = 0.0162 (gape area displayed (in mm^2)) + 0.178 (calls per 6 s) + 8.23 . Observed provisioning rates of cuckoo young are shown in blue, and of reed warbler broods in red. Each point represents mean gape area, number of calls per 6 s and feeds delivered per hour for chicks of different ages and/or broods of different sizes. Lines connect to points that lie above or below the plane. Cuckoos display a relatively small gape, but call at a relatively high rate, while the reverse is true for reed warbler broods. However, the provisioning rates of both fall close to the predicted plane, showing that reed warbler parents follow the same integration rule when feeding cuckoos and their own brood.

Methods

Measurement of nestling gape area, mass and call rate in the field. For all species, we used Vernier callipers to measure the distance between the proximal outer corners of the mouth flange to the nearest 0.05 mm (gape width), and the distance between the proximal corner and the tip (gape length) and calculated the product (gape area, assuming the usual fully open gape). For broods, total gape area displayed was the sum of the individual gapes. We found it impossible to handle cuckoo chicks older than day 11 and be certain that they would settle back in the nest. Therefore gape areas for cuckoo chicks on days 13 and 15 were estimated from the extrapolated polynomial regression of gape area on nestling weight, using published masses of common cuckoos in reed warbler nests at these ages³². We measured nestling mass in the field to the nearest 0.25 g using a 50 g Pesola spring balance, and in the laboratory to the nearest 0.005 g using a Sartorius electronic balance. We recorded the vocalizations emitted by unmanipulated cuckoos at different ages and reed warblers of different ages and in different brood sizes onto digital audiotape (DAT), using a Sony ECM-T6 tie-clip microphone attached to the side of the nest, and then recorded these calls onto a Macintosh computer. We scored the number of calls produced during the first 6 s of each nest visit (the shortest nest visit duration) using the sonograms produced by the application Canary 1.2.1.

Measurement of begging displays and food consumption in the laboratory. To standardize levels of food deprivation, we recorded the begging displays of chicks in the laboratory in June–July 1996–1998. Chicks were temporarily removed from the nest at a range of different ages (reed warblers day 2–8, cuckoos day 2–11) and in broods of one (reed warblers $n = 37$, cuckoos $n = 25$) or four ($n = 19$) chicks. They were kept warm⁵ and fed with Nectarblend rearing mix until they stopped begging. We considered this to be the moment of satiation. The broods of four chicks were induced to beg every 10 min for the next 110 min, and all chicks were induced to beg after 110 min. Repeated testing had no effect on begging behaviour^{5,17}. Begging was induced in the box apparatus described in ref. 5. Vocalizations were recorded as above. After begging at 110 min, we recorded how many small balls of Nectarblend all nestlings (except seven broods of four reed warblers) consumed until they stopped begging, and then returned them to their nests in the field. To standardize each test for comparison between species, we measured how many balls of food displaced 1 ml water.

Parental response to nestling visual and vocal signals. In June–July 1996–1998, we performed 111 treatments using 63 different pairs of parents when chicks were 6/7 days old. For each treatment we scored total feeds delivered per hour at the nest by both parents. Data were collected either by using a video camera ($n = 20$) or by observing the nest from a distance through binoculars ($n = 91$). Twenty-four pairs of parents, with brood sizes ranging from two to four chicks ($n = 10$ broods of 2; $n = 3$ broods of 3; $n = 11$ broods of 4) were each tested in three different treatments: playback of no vocalizations, playback of a single chick calling, and playback of four chicks calling. We rotated the sequence of testing between each pair of parents. Whether we used one randomly selected treatment per pair or the total number of treatments as independent data points, the outcome of our analysis was the same. Using each pair as a data point, there were significant independent effects of the number of gapes displayed ($t_{62} = 3.66$, $P = 0.0005$) and the number of chicks calling ($t_{62} = 4.36$, $P < 0.0001$) on parental provisioning rate.

Manipulation of brood size. We manipulated brood sizes so that they ranged from one to eight nestlings ($n = 59$). For broods of more than five chicks, old *Turdus* nests were placed over the top of the reed warbler nest and the enlarged brood was placed inside. Parents and chicks were given 1 h to adjust to the manipulation, which we have shown previously to be sufficient^{13,25}, and provisioning rate was measured for the next hour.

Playback manipulation. We broadcast the vocalizations of either a single reed warbler chick ($n = 28$) or a brood of four reed warblers ($n = 24$) through a speaker attached at the side of the nest at brood sizes ranging from two to five ($n = 52$). We played back recordings of 20 different single reed warbler chicks ($n = 14$ recorded in the laboratory, $n = 6$ recorded in the field) and 18 recordings of different broods of four reed warblers ($n = 8$ recorded in the laboratory, $n = 10$ recorded in the field). Details of playback technique and recordings of vocalizations are given elsewhere¹³.

Received 11 November 1998; accepted 13 January 1999.

1. Zuk, M., Ligon, J. D. & Thornhill, R. Effects of experimental manipulation of male secondary sex characters on female mate preference in red jungle fowl. *Anim. Behav.* **44**, 999–1006 (1992).
2. Marchetti, K. The evolution of multiple male traits in the yellow-browed leaf warbler. *Anim. Behav.* **55**, 361–372 (1998).
3. Scheffer, S. J., Uetz, G. W. & Stratton, G. E. Sexual selection, male morphology, and the efficacy of courtship signalling in two wolf spiders (Araneae: Lycosidae). *Behav. Ecol. Sociobiol.* **38**, 17–23 (1996).
4. Rowe, C. & Guilford, T. Hidden colour aversions in domestic chicks triggered by pyrazine odours of insect warning displays. *Nature* **383**, 520–522 (1996).
5. Kilner, R. M. & Davies, N. B. Nestling mouth colour: ecological correlates of a begging signal. *Anim. Behav.* **56**, 705–712 (1998).
6. Redondo, T. & Arias de Reyna, L. Locatability of begging calls in nestling altricial birds. *Anim. Behav.* **36**, 653–661 (1988).
7. Møller, A. P. & Pomiankowski, A. Why have birds got multiple sexual ornaments? *Behav. Ecol. Sociobiol.* **32**, 167–176 (1993).
8. Schluter, D. & Price, T. Honesty, perception and population divergence in sexually selected traits. *Proc. R. Soc. Lond. B* **253**, 117–122 (1993).
9. Iwasa, Y. & Pomiankowski, A. The evolution of mate preferences for multiple sexual ornaments. *Evolution* **48**, 853–867 (1994).
10. Johnstone, R. A. Honest advertisement of multiple qualities using multiple signals. *J. Theor. Biol.* **177**, 87–94 (1995).
11. Johnstone, R. A. Multiple displays in animal communication: 'backup signals' and 'multiple messages'. *Phil. Trans. R. Soc. Lond. B* **351**, 329–338 (1996).
12. Johnstone, R. A. in *Behavioural Ecology* (eds Krebs, J. R. & Davies, N. B.) 155–178 (Blackwell, Oxford, 1997).
13. Davies, N. B., Kilner, R. M. & Noble, D. G. Nestling cuckoos, *Cuculus canorus*, exploit hosts with begging calls that mimic a brood. *Proc. R. Soc. Lond. B* **265**, 673–678 (1998).
14. Dawkins, R. & Krebs, J. R. Arms races between and within species. *Proc. R. Soc. Lond. B* **205**, 489–511 (1979).
15. Dawkins, R. *The Selfish Gene* (Oxford Univ. Press, Oxford, 1989).
16. Kilner, R. & Johnstone, R. A. Begging the question: are offspring solicitation behaviours signals of need? *Trends Ecol. Evol.* **12**, 11–15 (1997).
17. Kilner, R. Mouth colour is a reliable signal of need in begging canary nestlings. *Proc. R. Soc. Lond. B* **264**, 963–968 (1997).
18. Price, K., Harvey, H. & Ydenberg, R. Begging tactics of nestling yellow-headed blackbirds, *Xanthocephalus xanthocephalus*. *Anim. Behav.* **51**, 421–435 (1996).
19. Trivers, R. Parent-offspring conflict. *Am. Zool.* **14**, 249–264 (1974).
20. Mock, D. W. & Parker, G. A. *The Evolution of Sibling Rivalry* (Oxford Univ. Press, Oxford, 1997).
21. Godfray, H. C. J. Signalling of need by offspring to their parents. *Nature* **352**, 328–330 (1991).
22. Godfray, H. C. J. Evolutionary theory of parent-offspring conflict. *Nature* **376**, 133–138 (1995).
23. Davies, N. B. & Brooke, M. de L. Cuckoos versus reed warblers: adaptations and counter adaptations. *Anim. Behav.* **36**, 262–284 (1988).
24. Jenner, E. Observations on the natural history of the cuckoo. *Phil. Trans. R. Soc. Lond. B* **78**, 219–237 (1788).
25. Brooke, M. de L. & Davies, N. B. Provisioning of nestling cuckoos *Cuculus canorus* by reed warbler *Acrocephalus scirpaceus* hosts. *Ibis* **131**, 250–256 (1989).
26. Cramp, S. *The Birds of the Western Palearctic* Vols IV, IX (Oxford Univ. Press, Oxford, 1995).
27. Noble, D. G., Davies, N. B., Hartley, I. R. & McRae, S. B. The red gape of the nestling cuckoo (*Cuculus canorus*) is not a supernormal stimulus for three common cuckoo hosts. *Behaviour* (submitted).
28. Guilford, T. & Dawkins, M. S. Receiver psychology and the evolution of animal signals. *Anim. Behav.* **42**, 1–14 (1991).
29. Ryan, M. J. in *Behavioural Ecology* (eds Krebs, J. R. & Davies, N. B.) 179–202 (Blackwell, Oxford, 1997).
30. Alcock, J. *Animal Behaviour* (Sinauer, Sunderland, 1998).
31. Harper, A. B. The evolution of begging: sibling competition and parent-offspring conflict. *Am. Nat.* **128**, 99–114 (1986).
32. Godfray, H. C. J. Signaling of need between parents and young: parent-offspring conflict and sibling rivalry. *Am. Nat.* **146**, 1–24 (1995).
33. Wyllie, I. *The Cuckoo* (Batsford, London, 1981).
34. Redondo, T. Exploitation of host mechanisms for parental care by avian brood parasites. *Etologia* **3**, 235–297 (1993).
35. Lichtenstein, G. & Sealy, S. G. Nestling competition, rather than supernormal stimulus, explains the success of parasitic brown-headed cowbird chicks in yellow warbler nests. *Proc. R. Soc. Lond. B* **265**, 249–254 (1998).

Acknowledgements. We thank NERC for funding the research; the National Trust for allowing us to work at Wicken Fen; C. Thorne and the Wicken Fen Group for research facilities; M. de L. Brooke for assistance in finding nests; M. de L. Brooke, F. Hunter, R. Johnstone and N. Langmore for comments on the manuscript; and B. Grenfell for producing Fig. 7. The experiments were done under licence from English Nature. R.M.K. was funded by a Junior Research Fellowship at Magdalene College, Cambridge.

Correspondence and requests for materials should be addressed to R.M.K. (e-mail: rmk1002@zoo.cam.ac.uk).

Aharonov–Bohm oscillations in carbon nanotubes

Adrian Bachtold*, Christoph Strunk*, Jean-Paul Salvetat†, Jean-Marc Bonard†, Laszlo Forró†, Thomas Nussbaumer* & Christian Schönenberger*

* Institut für Physik, Universität Basel, Klingelbergstrasse 82, CH-4056 Basel, Switzerland

† Institut de Génie Atomique, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland

When electrons pass through a cylindrical electrical conductor aligned in a magnetic field, their wave-like nature manifests itself as a periodic oscillation in the electrical resistance as a function of the enclosed magnetic flux¹. This phenomenon reflects the dependence of the phase of the electron wave on the magnetic field, known as the Aharonov–Bohm effect², which causes a phase difference, and hence interference, between partial waves encircling the conductor in opposite directions. Such oscillations have been observed in micrometre-sized thin-walled metallic cylinders^{3–5} and lithographically fabricated rings^{6–8}. Carbon nanotubes^{9,10} are composed of individual graphene sheets rolled into seamless hollow cylinders with diameters ranging from 1 nm to about 20 nm. They are able to act as conducting molecular wires^{11–18}, making them ideally suited for the investigation of quantum interference at the single-molecule level caused by the Aharonov–Bohm effect. Here we report magnetoresistance measurements on individual multi-walled nanotubes, which display pronounced resistance oscillations as a function of magnetic flux. We find that the oscillations are in good agreement with theoretical predictions for the Aharonov–Bohm effect in a hollow conductor with a diameter equal to that of the outermost shell of the nanotubes. In some nanotubes we also observe shorter-period oscillations, which might result from anisotropic electron currents caused by defects in the nanotube lattice.

In a diffusive and thin-walled metallic cylinder, a prominent periodic quantum correction to the resistance arises from the interference of closed electron trajectories that encircle the cylinder once. The phase difference $\Delta\phi$ between each such trajectory Γ and the time-reversed counter-propagating trajectory Γ' (Fig. 1a) is solely determined by the magnetic flux Φ enclosed: $\Delta\phi = 2\pi(2e/h)\Phi$, where e and h are the electron charge and Planck's constant, respectively. Consequently, the electrical resistance has an oscillating contribution with period $h/2e$, known as the Altshuler–Aronov–Spivak (AAS) effect¹. For zero magnetic flux, these interference terms add up constructively, increasing electron back-scattering and thereby the electrical resistance, an effect known as weak localization¹⁹. For a thin metallic film in a perpendicular magnetic field, the weak-localization resistance correction monotonically disappears in higher fields (negative magnetoresistance, MR). In contrast, for a cylinder in a parallel magnetic field, weak localization is periodically modulated with a magnetic-field period given by $\Delta B = (h/2e)/r^2\pi$, where r is the radius of the cylinder.

We have carried out electric-transport measurements on multi-walled carbon nanotubes (MWNTs) composed of multiple coaxial graphene cylinders, in a magnetic field parallel to the axis of the nanotubes. An example is shown in Fig. 1b. Coulomb blockade can strongly affect electrical transport in small structures, and we therefore use samples containing only a single contacted nanotube with low contact resistances (that is, not exceeding a few kilohms). Although this approach will not completely exclude Coulomb blockade effects, the different measurements and control experiments carried out indicate that Coulomb blockade is, at most, of minor importance. For example, we have never observed that

cooling leads to an increase in the resistance of a nanotube towards an isolating state, as would be expected for Coulomb blockade, and current–voltage characteristics were checked to be linear to at least 10 mV. We have also performed control experiments using an additional gate electrode. For electrostatic gating fields of up to 10^8 V m^{-1} , we observed resistance changes not exceeding 10%. Although the changes depend weakly on the electric field, we have not been able to gate the nanotube into an insulating state, as would occur if Coulomb blockade were important.

Figure 2 shows the measured electrical resistance (solid curves) of a MWNT as a function of magnetic field B for different temperatures. A large modulation of the resistance ($\sim 30\%$) is observed, comparable in size to the quantum resistance $h/2e^2$. Starting from $B = 0$, the resistance decreases with increasing field. Previous studies in perpendicular magnetic fields observed, a similar decrease in MR which could be interpreted within the framework of weak localization^{11,17}. It thus seems reasonable to assume that the same physics is responsible for the resistance change seen here. At higher magnetic fields, a second resistance peak develops at $\pm 8.8 \text{ T}$. This second peak is assigned to the reappearance of enhanced back-scattering due to the AAS effect. As expected, a similar peak is not observed in perpendicular fields.

To provide further evidence for the resistance oscillations constituting AAS effects, we compare our measurements with theoretical predictions for the AAS effect in cylinders^{1,4}. We first determine the

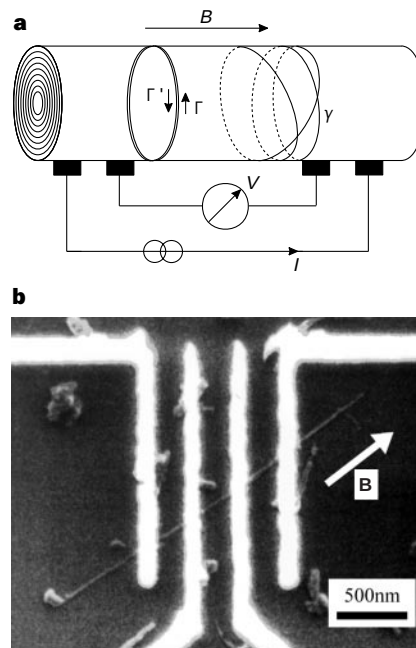


Figure 1 A single multi-walled carbon nanotube electrically contacted with four electrodes. **a**, Diagram of a multi-walled nanotube (MWNT), which is composed of a series of coaxial graphene cylinders. The electrical resistances of single MWNTs are measured via four electrical contacts (imposed current I and measured voltage V) in a magnetic field B which is aligned parallel to the tube axis to within a few degrees. A periodic magnetoresistance is expected to originate from quantum interference of counter-propagating closed diffusive electron trajectories, like those denoted by Γ and Γ' or γ . **b**, Scanning-electron microscopy (SEM) image of a MWNT contacted by four Au fingers; the direction of B is shown by a white arrow. The MWNTs are produced by arc-discharge evaporation²⁷. After purification²⁸, they are adsorbed from a dispersion in chloroform onto an oxidized Si wafer. An electrode structure, consisting of four Au fingers which are 70 nm thick, $\sim 100 \text{ nm}$ wide, $2 \mu\text{m}$ long, and separated by 350 nm (centre-to-centre), is then patterned onto the nanotubes using electron-beam lithography. As a SEM image can only give a qualitative estimate for the diameter of the nanotube, atomic force microscopy is used to determine the diameter of an electrically measured MWNT.

effective (that is, the mean current-carrying) cylinder radius r_e from the peak position at a magnetic field of 8.8 T. As this field ought to correspond to a magnetic flux of $h/2e$ threading the cross-section πr_e^2 , a radius of $r_e = 8.6 \pm 0.1$ nm is inferred. This value is in excellent agreement with the outer radius $r_0 = 8 \pm 0.8$ nm of the nanotube as determined by atomic force microscopy and not, as one might have expected, with the mean radius of all graphene cylinders of this MWNT (expected to be ≈ 5 nm).

For comparison with theory, we first consider a single graphene cylinder with known radius r_e . Then, three parameters appear in the AAS theory: the misalignment angle Θ between the magnetic field direction relative to the cylinder axis, the length L of the current-carrying part of the cylinder between the contacts, and the temperature-dependent phase-coherence length l_ϕ . We note that a value of Θ of a few degrees can account for the observed reduced amplitude of the resistance peaks at ± 8.8 T (ref. 4). Matching the theory to the measurement at $T = 1.8$ K, we obtain $l_\phi = 54$ nm, $L = 170$ nm and $\Theta = 4.4^\circ$. These values of Θ and L agree well with the geometry of this sample. Keeping Θ and L constant, a similar procedure is applied to the other MR curves. As illustrated in Fig. 2, we find good agreement between theory (dotted lines) and experiment (solid lines). This agreement is obtained while assuming that only one graphene cylinder contributes to the conductance. We may relax this assumption by considering a set of concentric cylinders which carry the electrical current equally. Under the constraints given by the measured outer diameter (for possible radii r) and the distance between contacts (for L), we find that at most two outer shells can contribute; this is much less than the number of shells that make up an actual MWNT (in our case, ~ 20 are expected).

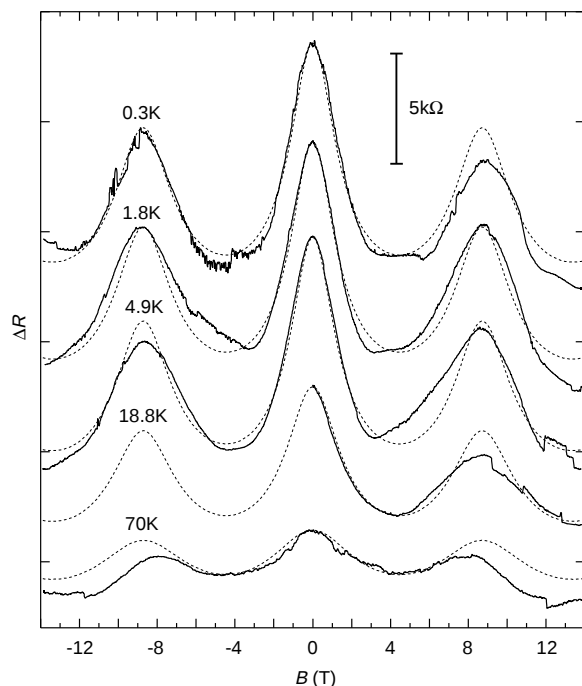


Figure 2 Magnetoresistance $R(B)$ at five different temperatures T measured for a MWNT in a parallel magnetic field B (solid curves). The dotted curves correspond to theoretical predictions for the quantum correction to the resistance of a thin-walled cylindrical conductor using the parameters: $r_e = 8.6$ nm, $\theta = 4.4^\circ$ and $L = 170$ nm for the effective cylinder radius, the angle between magnetic field and cylinder axis, and the length of the cylinder between the electric contacts, respectively. All curves are vertically offset for clarity. At zero field the absolute values for the resistances are (kΩ): 30.6, 30.1, 29.8, 25 and 21.4 from top to bottom ($T = 0.3, \dots, 70$ K). Non-reproducible temporal resistance jumps are observed in the experiment, possibly caused by temporal changes in the electric contacts, defects or impurities in the nanotubes.

We therefore conclude that quantum-interference corrections to the resistance can account for the measured MR, if and only if the electrical current is assumed to flow in one (or two) of the outermost shells. This finding may be understood as follows: by virtue of the fabrication technique, contact is made to the outer surface of the nanotube, so that current is injected into the outermost conducting graphene cylinder. As nanotubes can be conducting or insulating^{20–24}, the outermost conducting tube may be followed by an insulating one, restricting the conduction to one cylinder. It is also possible that the anisotropy of the conductivity, which resembles graphite²⁵, is enough to keep the current confined to one cylinder, in particular at low temperatures where inter-shell hopping is suppressed. Owing to the small current-carrying cross-section, a remarkably low resistivity is obtained. We estimate this resistivity to be $\leq 10 \mu\Omega$ cm at room temperature, which is comparable to that of metals with good conductivity.

In addition to the MR shown in Fig. 2, we have also observed an oscillatory feature on several other MWNTs (we believe that this latter feature has not been previously observed). Figure 3 shows such a measurement for two temperatures for a nanotube of outer radius $r_0 = 6.5 \pm 0.5$ nm determined by atomic force microscopy. The most striking and immediately visible difference from the previously mentioned sample is a superimposed short-period oscillation (see arrows). In agreement with the MR of the previous sample, the resistance (neglecting the rapid oscillation) at first decreases with increasing magnetic field. Then, it develops a minimum at ~ 9 T and increases again, possibly towards a second maximum, which is beyond the experimental field range and is expected at ~ 18 T from the measured size of the tube. The linear

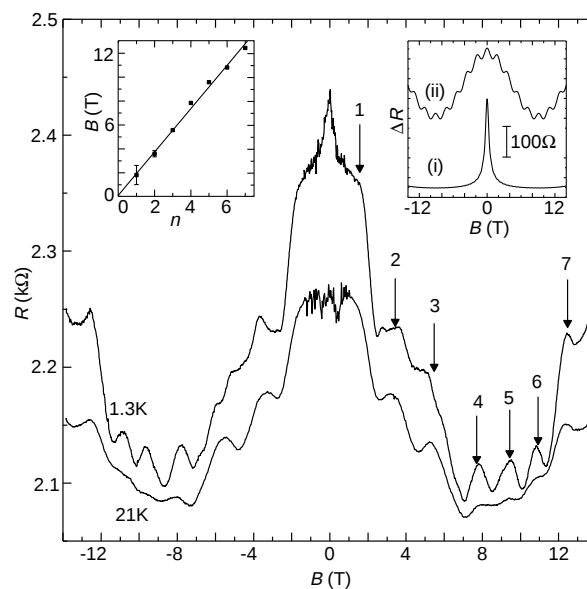


Figure 3 As Fig. 2 but for two different temperatures T on another MWNT. The general decrease of the resistance from 0 T to ~ 9 T, and the subsequent increase to a possible second maximum (beyond the magnetic-field range of the experiment), is assigned to the fundamental oscillation of the AAS effect. In addition, a superimposed oscillation with a much smaller period is clearly visible. Positions of maxima in resistance for this oscillation are labelled by arrows and plotted in the left inset, which shows that the oscillation is periodic with period $\Delta B_{\text{fast}} = 1.8$ T. The peak positions agree for both temperatures. This oscillation is interpreted to originate from electron interference of preferentially selected higher-order loops which wind several times around the nanotube (see also Fig. 1a). Right inset, theoretically predicted MR for (i) the 'full' AAS effect (that is, adding up corrections from loops with all winding numbers) and (ii) the case when only trajectories with winding number 1 and 10 are considered. For both cases, we require the magnitude of the resistance correction to match the experimental data. The strong temporal resistance fluctuations, which appear near zero field, are not yet understood, but disappear for larger fields $|B| \geq 1.5$ T.

dependence of the peak positions of the short-period oscillation on magnetic field (Fig. 3 left inset) demonstrates that a periodic oscillation is observed with period $\Delta B_{\text{fast}} = 1.8 \text{ T}$. ΔB_{fast} cannot correspond to the fundamental $h/2e$ period of the AAS effect, as this condition would require a cylinder with an effective radius of $r_c = 19 \text{ nm}$, which is significantly larger than the outer radius for this nanotube. As the two peaks denoted by 3 and 7 appear approximately symmetrical with respect to the gross-minimum of the MR curve, the minimum is assigned to coincide with peak number 5. The fundamental AAS period then comprises $n = 10$ 'fast' oscillations.

Being convinced about the assignment for the fundamental AAS period, we ask where the 'fast' magneto-oscillation may come from. Closed electron trajectories that encircle the tube once (winding number 1) give rise to the fundamental magnetic flux-period $h/2e$. Trajectories with a larger winding number $n > 1$ (see, for example, the $n = 3$ loop denoted by γ in Fig. 1a) contribute to the resistance with an oscillating term given by $K_0(n2\pi r/l_0)\cos(2\pi n\Phi/(h/2e))$; the amplitude is determined by the modified Bessel function $K_0(x)$ (refs 1, 4). Though such a contribution has a flux-period which is n -times smaller than $h/2e$, the present theory cannot account for the experimental result. The reason is that the sum over all terms $n = 1, \dots, \infty$ of this Fourier series describes a smooth periodic function with period $h/2e$. Higher-order oscillations, as observed, can only show up if trajectories with specific winding numbers are preferentially selected. In order to allow a qualitative comparison, we display in the right inset of Fig. 3 the theoretical resistance dependence for two cases: (i) the contributions from all winding numbers are summed according to theory, (ii) only the $n = 1$ and $n = 10$ terms are retained. The latter resembles the measurement much more closely. From the condition that the amplitude of the $n = 1$ and $n = 10$ terms should agree with the measurement, the phase-coherence length is estimated to be $l_\phi \approx 250 \text{ nm}$, which is much larger (by a factor of 6.5) than the circumference $2\pi r_c$, in contrast to the previous example for which $l_\phi \approx 2\pi r_c$. This difference may be the reason why specific higher-order terms do not appear in the measurement of Fig. 2. We note that taking only two terms with distinct winding numbers to contribute to the resistance correction is justified as a first guess, but certainly oversimplifies the problem. In general, there is a distribution of weights for the higher-order terms which is specific for each nanotube, and is observed to differ from the AAS prediction.

Although we have an understanding of the long-period oscillations, the appearance of higher-order terms is at present unexplained. One possibility is that the short-period oscillations are due to chiral currents in strained nanotubes²⁶. Chirality is a well established structural property of carbon nanotubes. In principle, graphitic tubules have isotropic conductivity; but when they are mechanically stretched owing to the interaction with the substrate, or owing to thermal contracting, the 'honeycomb' lattice can be distorted resulting in anisotropic, chiral current. Chirality can select particular winding numbers which give the higher-order term superimposed on the long-period oscillations observed in Fig. 3. In an MWNT the chirality can change from layer to layer, and the period of short oscillations can vary from tube to tube and from sample to sample. This fine structure in the AAS oscillations is closely related to the special structure of the carbon nanotubes, and further experimental and theoretical studies are needed for its detailed understanding. □

Received 25 August; accepted 8 December 1998.

- Altshuler, B. L., Aronov, A. G. & Spivak, B. Z. The Aharonov–Bohm effect in disordered conductors. *Pis'ma Zh. Eksp. Teor. Fiz.* **33**, 101–103 (1981) [*JETP Lett.* **33**, 94–97 (1981)].
- Aharonov, Y. & Bohm, D. Significance of electromagnetic potentials in the quantum theory. *Phys. Rev.* **115**, 485–491 (1959).
- Sharvin, D. Y. & Sharvin, Y. V. Magnetic-flux quantization in a cylindrical film of a normal metal. *Pis'ma Zh. Eksp. Teor. Fiz.* **34**, 285–288 (1981) [*JETP Lett.* **34**, 272–275 (1982)].
- Aronov, A. G. & Sharvin, Y. V. Magnetic-flux effects in disordered conductors. *Rev. Mod. Phys.* **59**, 755–779 (1987).

- Gijs, M., Van Haesendonck, C. & Bruynseraede, Y. Resistance oscillations and electron localization in cylindrical Mg films. *Phys. Rev. Lett.* **52**, 2069–2072 (1984).
- Pannetier, B., Chaussy, J., Rammal, R. & Gandit, P. Magnetic flux quantization in the weak-localization regime of a nonsuperconducting metal. *Phys. Rev. Lett.* **53**, 718–721 (1984).
- Webb, R., Washburn, S., Umbach, C. & Laibowitz, D. C. Observation of h/e Aharonov–Bohm oscillations in normal-metal rings. *Phys. Rev. Lett.* **54**, 2696–2699 (1985).
- Chandrasekhar, V., Rooks, M. J., Wind, S. & Prober, D. E. Observation of Aharonov–Bohm electron interference effect with period h/e and $h/2e$ in individual micro-sized, normal-metal rings. *Phys. Rev. Lett.* **55**, 1610–1613 (1985).
- Ijima, S. Helical microtubules of graphitic carbon. *Nature* **354**, 56–58 (1991).
- Dresselhaus, M. S., Dresselhaus, G. & Eklund, P. C. *Science of Fullerenes and Carbon Nanotubes* (Academic, San Diego, 1986).
- Langer, L. *et al.* Quantum transport in multiwalled carbon nanotubes. *Phys. Rev. Lett.* **76**, 479–482 (1996).
- Dai, H., Wong, E. W. & Lieber, C. M. Probing electrical transport in nanomaterials: conductivity of individual carbon nanotubes. *Science* **272**, 523–526 (1996).
- Ebbesen, T. W. *et al.* Electrical conductivity of individual carbon nanotubes. *Nature* **382**, 54–56 (1996).
- Tans, S. J. *et al.* Individual single-wall carbon nanotubes as quantum wires. *Nature* **386**, 474–477 (1997).
- Bockrath, M. *et al.* Single-electron transport in ropes of carbon nanotubes. *Science* **275**, 1922–1925 (1997).
- Bachtold, A. *et al.* Contacting carbon nanotubes selectively with low-ohmic contacts for four-probe electric measurements. *Appl. Phys. Lett.* **73**, 274–276 (1998).
- Bachtold, A., Strunk, C., Schönenberger, C., Salvetat, J.-P. & Forró, L. In *Proc. XIIIth Int. Winter School on Electronic Properties of Novel Materials* (eds Kuzmany, H., Fink, J., Mehring, M. & Roth, S.) 65–68 (AIP, New York, 1998).
- Frank, S., Poncharal, P., Wang, Z. L. & Heer, W. A. Carbon nanotube quantum resistors. *Science* **280**, 1744–1746 (1998).
- Begmann, G. Weak localizations in thin films. *Phys. Rep.* **107**, 1–56 (1984).
- Hamada, N., Sawada, D.-I. & Oshiyama, A. New one-dimensional conductors: graphite microtubules. *Phys. Rev. Lett.* **68**, 1579–1581 (1992).
- Saito, R., Fujita, M., Dresselhaus, G. & Dresselhaus, M. S. Electronic structure of chiral graphene tubules. *Appl. Phys. Lett.* **60**, 2204–2206 (1992).
- Mintmire, J. W., Dunlap, B. I. & White, C. T. Are fullerene tubules metallic? *Phys. Rev. Lett.* **68**, 631–634 (1992).
- Wildöer, J. W. G., Venema, L. C., Rinzler, A. G., Smalley, R. E. & Dekker, C. Electronic structure of atomically resolved carbon nanotubes. *Nature* **391**, 59–62 (1998).
- Odom, T. W., Huang, J.-L., Kim, P. & Lieber, C. M. Atomic structure and electronic properties of single-walled carbon nanotubes. *Nature* **391**, 62–64 (1998).
- Wallace, P. R. The band theory of graphite. *Phys. Rev.* **71**, 622–634 (1947).
- Miyamoto, Y. Mechanically stretched carbon nanotubes: Induction of chiral current. *Phys. Rev. B* **54**, R11149–R11152 (1996).
- Ebbesen, T. W. & Ajayan, P. M. Large-scale synthesis of carbon nanotubes. *Nature* **358**, 220–222 (1992).
- Bonard, J. M. *et al.* Purification and size-selection of carbon nanotubes. *Adv. Mater.* **9**, 827–831 (1997).

Acknowledgements. We thank H.-W. Fink, M. Henny, T. Hoss, M. Krüger, C. Terrier and V. Thommen for contributions. This work was supported by the Swiss National Science Foundation.

Correspondence and requests for materials should be addressed to C.S. (e-mail: Schonenberg@ubacul.unibas.ch).

Spontaneous chaotic granular mixing

Troy Shinbrot, Albert Alexander & Fernando J. Muzzio

Pharmaceutical Engineering Program, Department of Chemical and Biochemical Engineering, Rutgers University, Piscataway, New Jersey 08854, USA

There are several types of instabilities in fluid mechanics that lead to spontaneous chaotic mixing and intricate patterns. Classical examples include the Kelvin–Helmholtz instability^{1,2} in shear layers, the instability of Taylor–Couette flow between rotating cylinders^{3,4} and the Rayleigh–Bénard instability in thermal convection⁵. More recently, a variety of two- and three-dimensional chaotic mixing phenomena have been observed in other geometries^{6–9}. Mixing in granular flows^{10,11}, unlike that in stirred fluids, is thought to be diffusive—although periodic forcing has been used to enhance granular mixing^{12,13}, spontaneous chaotic granular mixing has not previously been reported. Here we report the observation of chaotic granular mixing patterns in simple cylindrical tumblers partially filled with fine grains. The patterns form spontaneously when sufficiently fine grains ($\approx 300 \mu\text{m}$ diameter) are blended. We identify the mechanism by which the chaotic patterns are produced: a periodic stick–slip behaviour occurs in the shear layer separating static and flowing regions of grains. This causes weakly cohesive grains to mix at rates over-

whelmingly exceeding those achievable for previously studied^{11,14} freely flowing grains.

Motivated by the considerable industrial importance of granular processing—accounting for an estimated thousand million tonnes of production annually¹⁵—the study of granular tumblers has a long history. Research spanning three decades has included radial and axial segregation^{16,17}, two-dimensional mixing^{11,14}, fingering¹⁸, and transverse wave dynamics¹⁰. Magnetic resonance imaging¹⁹ has confirmed that the time-averaged velocity field within a cylindrical tumbler operating in the “rolling” regime¹⁵ consists of a flat granular layer flowing downhill above a region of returning grains that rotate with the tumbler as a solid body²⁰. Experimental²¹ and computational²² evidence indicate that some complex, industrially relevant, mixers exhibit similar phenomenology.

We emphasize that there is necessarily a shear layer between the flowing and the solid-body regions, and it is known that granular shear layers can exhibit periodic stick–slip behaviour²³. This stick–slip motion consists of slow creeping flow alternating with rapid slip events, and is most common in fine or cohesive powders widely used in pharmaceutical, ceramic and food processing. Periodic oscillations in simple flows are known to produce chaotic mixing⁹, and here we report observations of intricate mixing patterns in granular tumblers which we duplicate using a periodically disturbed and chaotic flow model. We produce experimental mixing patterns (Fig. 1a) in small tumblers partially filled with fine glass beads. Companion numerical results (Fig. 1b) are generated using a periodically forced continuum model of the tumbler, which we describe below. In both experiment and model, the grains are identical except for colour. Figure 1 inset shows an enlarged view of the fine striations that persist down to the grain size of the blend.

The experiments are performed using cylindrical metal tumblers

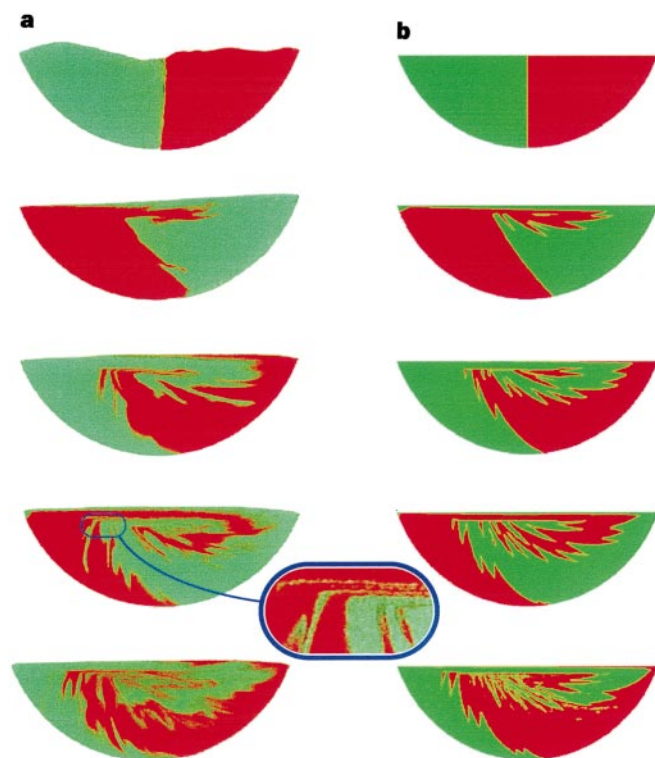


Figure 1 Comparison between experiments (a) and simulation (b) of mixing in rotating cylindrical tumblers one-third filled with identical red and green particles. From top to bottom, the mixing pattern is shown after 0, 0.25, 0.5, 0.75 and 1 revolution of the tumbler. Inset, striated structure that persists down to the grain size of the material.

with length 14 cm and diameter $D = 10$ cm, loaded with small glass beads (Potters Industries, Carlstadt, New Jersey) to a depth $D/3$. Initially, green beads are placed upstream of the midplane and red beads downstream. All beads are spherical, sieved to have diameters between 63 and 180 μm . The tumblers are rotated at 8 ± 0.2 revolutions per min under stepper control for the prescribed duration, after which the tumbler is stopped and the mixture is ‘frozen’ by infiltrating it with methacrylate copolymer (Cheesebrough-Ponds, Greenwich, Connecticut). The polymer is allowed to set, and the solidified mixture is sliced open to reveal the internal mixing patterns. This freezing and slicing technique permits direct and undisturbed observation of the mixture structure within the blender¹². Companion freezing and slicing experiments in three-dimensional, “double-cone” tumblers of the sort widely used industrially reveal similar internal mixing structures. Near the walls of the cylindrical blender (and throughout thin, disk-like tumblers), these intricate patterns are suppressed (Fig. 3 left inset), and we have found that they vanish altogether in experiments using particles larger than about 300 μm diameter. This may explain why these patterns have not previously been reported.

The data in Fig. 1 suggest that, in a similar fashion to chaotic fluid flows, the granular mixing process is driven by an iterative, multiplicative, orientation-preserving mechanism. As portions of the interface between red and green particles reach a given position, they rapidly acquire an invariant orientation. This is a consequence of the fundamental “asymptotic directionality” property that is common to all periodic area-preserving chaotic flows^{9,24}. As a result of this property, a self-similar structure emerges: as time progresses, new features are incorporated within the existing skeleton previously established. During granular tumbling, each time material enters the flowing layer, it is stretched and folded by a stick–slip behaviour (which we define below) into jagged, ‘shark’s tooth’, patterns. In successive circuits through the layer, the procedure is repeated, and greater detail is added within each ‘shark’s tooth’. This iterative process causes the interface between constituents to grow exponentially rapidly; it also causes the length scale of the unmixed regions to decrease exponentially rapidly down to the diffusional limit in fluid flows, and down to the grain size for granular systems.

Two-dimensional model simulations are performed as follows. As we have indicated, the crucial physical phenomenon to be captured is a periodic stick–slip in the shear band between the flowing layer above and the solid-body region below, because chaos is precluded in two-dimensional deterministic flows without the introduction of such a time-dependence. In Fig. 2a, we show a schematic diagram of the model used for this stick–slip behaviour. Material in the presumed shear band (shown cross-hatched) alternately slowly creeps downstream and rapidly slips; as slip occurs, the interface between the flowing and the solid-body regions propagates upstream. Spontaneous surface disturbances are visually observed to propagate downstream, beginning near the upstream end of the tumbler (where compressive stress is least). These disturbances are seen in our tumbler for at least the range in speed 1–10 revolutions per min provided that small ($\leq 300 \mu\text{m}$) grains are used.

We therefore adopt the following procedure to simulate the granular flow. This is neither the only nor necessarily the most accurate model for the observed phenomenon; we choose this model because of its relative simplicity and its agreement with outward experimental observations. Initially, the flowing layer is taken to be separated from the solid-body region beneath by a symmetric parabola;

$$z = f(x) = \frac{D - 2F + d}{F(D - F)} x^2 + [F - d - (D/2)] \quad (1)$$

where, as shown in Fig. 2a, F is the linear fill level, d is the maximum thickness of the flowing layer, and D is the tumbler diameter. Particles beneath the parabola rotate with the angular speed, ω , of the tumbler, while particles above the parabola travel parallel to the

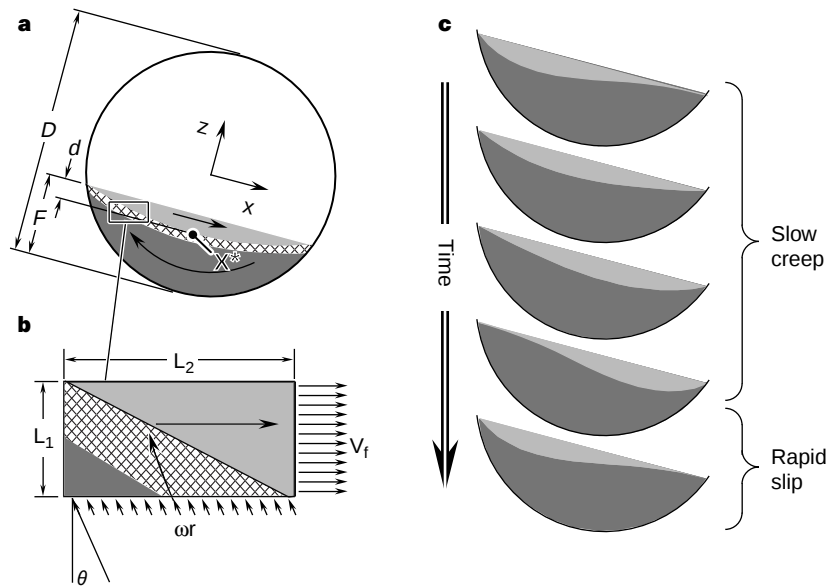


Figure 2 The continuum model. **a**, Geometry of the time-averaged model flow, with the flowing layer (light grey) above the solid-body region (dark grey). The shear band between these regions (cross-hatched) is considered to be unstable to slip events. **b**, Evolution of the flowing layer during periodic slow creep and

rapid slip epochs. **c**, To conserve mass, the speed of particles, V_f , leaving an interfacial region to the right equals the speed of particles entering the region from below, ωr , multiplied by $\cos(\theta)L_2/L_1$.

flat surface with a speed, V_f , defined to preserve the volume of material entering the flowing layer. Neglecting dilation, this is done by compensating for both the angle of incidence, θ , of material incoming from the solid-body region and the angle of the interface, $\tan^{-1}(L_1/L_2)$, between the solid-body and flowing regions, as sketched in Fig. 2b and described in the figure legend. Explicitly, it can be shown that to conserve mass, the speed in the flowing layer must obey the relation: $V_f = \omega r \cos(\theta)(L_2/L_1) = \omega/2a$, where r is the radial location of the particle at the point where it enters the flowing layer, and $a = (D - 2F + d)/[F(D - F)]$ is the coefficient of the parabolic interface, equation (1).

Although this flow agrees¹⁴ with time-averaged experimental data²¹ for large freely flowing grains, it does not predict the mixing patterns observed for small cohesive particles. So once per unit time, T , the layer is permitted to slip rapidly near the upstream end of the free surface, and thereafter to slowly creep downstream before repeating the cycle. In detail, $z = f[x + A(t)f(x)]$, where $A(t)$ is the sawtooth function:

$$A(t) = A_0 \cdot \begin{cases} 1 - (2t/\epsilon) & \text{if } 0 < (t \bmod T) \leq \epsilon \\ (2(t - \epsilon)/(T - \epsilon)) - 1 & \text{if } \epsilon < (t \bmod T) \leq T \end{cases} \quad (2)$$

Here t is time, A_0 is a constant amplitude and ϵ is an eccentricity, $\epsilon \ll T$, defining how rapidly the slip part of the cycle occurs. Equations (1) and (2) conserve area in the flowing layer to a close approximation²⁵ and guarantee that the layer thickness vanishes at the up- and down-stream extremes of the flowing surface.

This model produces the behaviour shown in Fig. 2c. The flowing layer creeps slowly downhill, and then rapidly slips to resume its initial state. In the results described here, we use $D = 2$, $d = 0.1$, $F = 0.33$, $A = 5$, $\epsilon = 20$ and $T = 1,000$ computational units, and we integrate particle trajectories using the Euler method with 10^4 timesteps per container revolution. Following recent studies in fluid mixing^{7,26}, we advect an interface dividing the filled volume in half, and we add new points to the interface whenever existing points become separated by a prescribed resolution, $0.02D$, except in the neighbourhood of a fixed point, X^* , which necessarily²⁶ occurs on the interface between the left-going solid-body region and the right-going flowing layer. For any fill level $F < 0.5D + d$, the interface

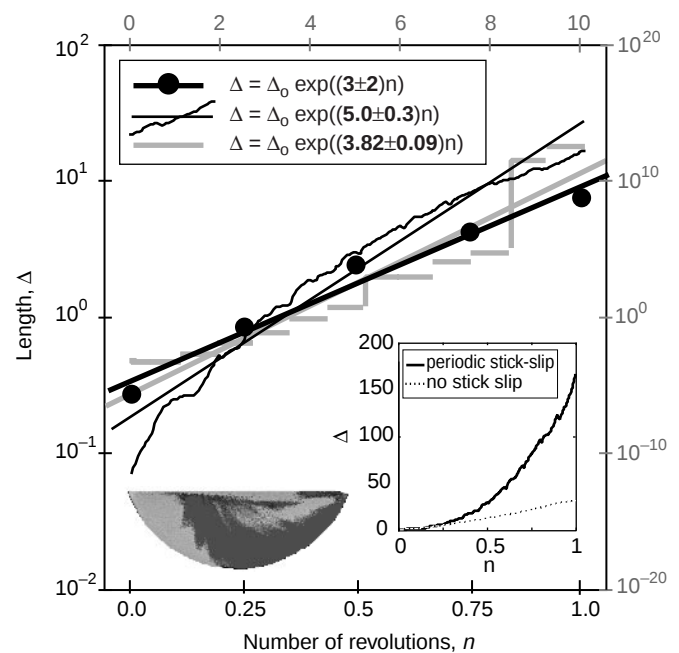


Figure 3 Stretching of experimental and numerical interfaces versus the number of tumbler revolutions. Experimental data (filled circles) are obtained by calculating the numbers of pixels separating the green and red areas in the photographs shown in Fig. 1a. Numerical data (thin curve) show stretching of the same interface in the computational model described in the text. The scales for both of these data sets are in black, on the bottom and the left of the plot. The bottom right inset shows a comparison on a linear scale between the stretching of the same interface with (solid curve) and without (broken curve) stick-slip. Grey data in the main plot are typical longer-time stretching results from numerical simulations extending up to 10 revolutions. The scales for these data are in grey, at the top and the right of the plot; the aspect ratio of vertical to horizontal scales are identical for all plots. Best fits to each data set are shown as lines, with slopes indicated in the key. Bottom left inset shows a suppressed mixing pattern visible through a side-wall of a cylinder after one revolution (dark grey indicates red particles, light grey indicates green).

winds infinitely rapidly around X^* , so to prevent computation divergence, interpolation is interrupted whenever an interfacial element cycles around X^* . This has the physical interpretation that material granularity causes particles near X^* to simply rotate in place^{27,28}.

Results from both model and experiments are summarized in Fig. 3. In the main plot, we show three measures of chaotic stretching. First, as filled circles, we plot the number of pixels, Δ , (normalized so that $\Delta = 1$ across the vessel diameter) bordering green and red regions in the photographs shown in Fig. 1a as a function of the number of rotations, n , of the experimental cylinder. Second, as a thin line, we plot the calculated length (again normalized to the vessel diameter) of the same interface from the model whose results are summarized in Fig. 1b. Each of these measures extends from 0 to 1 complete revolution of the tumbler. Third, we compute a longer-time stretching rate²⁹ by evolving a very short line segment (length $\Delta_0 = 0.0005D$) 10 revolutions forward in time, rescaling the length to Δ_0 whenever it grows above $0.005D$. In Fig. 3, we show in grey the growth starting from an initial condition centred on the line of symmetry through the granular bed at a radius $r = 0.375D$. This initial condition is used as an illustration because it produces a stretching rate (that is, the slope, K , of $\log(\Delta)$ versus n) close to the average, $\langle K \rangle = 3.7$ revolutions⁻¹, that one obtains from a uniform distribution of initial conditions along the symmetry line. K ranges from 0 to 8.7, with the smallest values near X^* , where the periodic perturbation traps particles, and near the cylinder walls, where (for low fill) particles seldom reach the flowing layer.

These three data sources are each fitted with an exponential function, giving stretching rates of 3 ± 2 , 5 ± 0.3 and 3.82 ± 0.09 revolutions⁻¹, respectively. In each case the stretching rate is unambiguously positive, consistent with exponential stretching over several orders of magnitude in only one or two tumbler revolutions. For comparison, in Fig. 3 bottom right inset, we plot on a linear scale the stretching of the same interface in the model with (solid curve) and without (broken curve) periodic stick-slip. (Behaviour without stick-slip corresponds to $A = 0$ in equation (2).) As anticipated, in a single revolution the periodically driven case exhibits strongly nonlinear interfacial stretching over two orders of magnitude, whereas the unforced case grows linearly with time, accounting for only one order of magnitude of interfacial stretching.

Evidently fine powders, like fluids, can spontaneously mix chaotically, at rates overwhelmingly exceeding those possible by the mechanisms expected from studies of freely flowing grains. Neither the spontaneous occurrence of the chaotic patterns studied here, nor the potential for exponential mixing rates, has been reported previously. From the fundamental standpoint, it seems clear that we still have much to learn about granular dynamics, as evidenced by the apparent connection between microscopic dynamics (which initiate stick-slip behaviour) and macroscopic transport (manifested by the intricate mixing patterns that we have reported). From the practical standpoint, the results of mixing studies of coarse grains clearly cannot be applied to the blending of fine or cohesive powders, which are widely used industrially. The challenge for the future will be to understand, from an applied perspective, how to exploit the connection between micro- and macro-scale granular dynamics to obtain improvements in mixing times which amount to tens of orders of magnitude. □

Received 9 October 1998; accepted 12 January 1999.

1. Thorpe, S. A. Experiments on the instability of stratified shear flows: miscible fluids. *J. Fluid Mech.* **46**, 299–319 (1971).
2. Van Dyke, M. (ed.) *An Album of Fluid Motion* (Parabolic, Stanford, California, 1982).
3. Taylor, G. I. Stability of a viscous liquid contained between two rotating cylinders. *Phil. Trans. R. Soc. Lond. A* **223**, 289–343 (1923).
4. Andereck, C. D., Liu, S. S. & Swinney, H. L. Flow regimes in a circular Couette system with independently rotating cylinders. *J. Fluid Mech.* **164**, 155–183 (1986).
5. Solomon, T. H. & Gollub, J. P. Chaotic particle transport in time-dependent Rayleigh-Bénard convection. *Phys. Rev. A* **38**, 6280–6284 (1988).
6. Aref, H. Stirring by chaotic advection. *J. Fluid Mech.* **143**, 1–21 (1984).

7. Alvarez, M. M., Muzzio, F. J., Cerbelli, S., Androver, A. & Giona, M. Self-similar spatio-temporal structure of material filaments in chaotic flows. *Phys. Rev. Lett.* **81**, 3395–3398 (1998).
8. Lamberto, D. J. & Muzzio, F. J. Experimental and computational investigation of the laminar flow structure in a stirred tank. *Chem. Eng. Sci.* (in the press).
9. Hobbs, D. M., Alvarez, M. M. & Muzzio, F. J. Mixing in globally chaotic flows: a self-similar process. *Fractals* **5**, 395–428 (1997).
10. Choo, K., Molteni, T. C. A. & Morris, S. W. Traveling granular segregation patterns in a long drum mixer. *Phys. Rev. Lett.* **79**, 2975–2977 (1997).
11. Metcalf, G., Shinbrot, T., McCarthy, J. J. & Ottino, J. M. Avalanche mixing of granular solids. *Nature* **374**, 39–41 (1995).
12. Wightman, C., Mort, P. R., Muzzio, F. J., Riman, R. E. & Gleason, E. K. The structure of mixtures of particles generated by time-dependent flows. *Powder Technol.* **84**, 231–240 (1995).
13. Brone, D. et al. Using flow perturbations to enhance mixing of dry powders in V-blenders. *Powder Technol.* **91**, 165–172 (1997).
14. Khakhar, D. V., McCarthy, J. J., Shinbrot, T. & Ottino, J. M. Transverse flow and mixing of granular materials in a rotating cylinder. *Phys. Fluids* **9**, 31–43 (1997).
15. Van Cleef, J. Powder technology. *Am. Sci.* **79**, 304–315 (1991).
16. Gupta, S. D., Khakhar, D. V. & Bhatia, S. K. Axial segregation of particles in a horizontal rotating cylinder. *Chem. Eng. Sci.* **46**, 1513–1525 (1991).
17. Hill, K. M., Caprihan, A. & Kakaliotis, J. Bulk segregation in rotated granular material measured by magnetic resonance imaging. *Phys. Rev. Lett.* **78**, 50–53 (1997).
18. Pouliquen, O., Delour, J. & Savage, S. B. Fingering in granular flows. *Nature* **386**, 816–817 (1997).
19. Nakagawa, M., Altobelli, S. A., Caprihan, A., Fukushima, E. & Jeong, E. K. Non-invasive measurements of granular flows by magnetic resonance imaging. *Exp. Fluids* **16**, 54–60 (1993).
20. Campbell, H. & Bauer, W. C. Cause and cure of demixing in solid-solid mixers. *Chem. Eng.* 179–185 (1966).
21. Chester, A. W. et al. Mixing dynamics in catalyst impregnation in twin-cone blenders. *Powder Technol.* (in the press).
22. Moakher, M., Shinbrot, T. & Muzzio, F. J. Experimentally validated computations of flow, mixing and segregation of non-cohesive grains in 3D tumbling blenders. (preprint, Rutgers Univ., Piscataway, New Jersey 08854 USA, 1998).
23. Nasono, S., Kudrolli, A. & Gollub, J. P. Friction in granular layers: hysteresis and precursors. *Phys. Rev. Lett.* **79**, 949–952 (1997).
24. Alvarez, M. M., Muzzio, F. J., Cerbelli, S., Androver, A. & Giona, M. Self-similar spatiotemporal structure of intermaterial boundaries in chaotic flows. *Phys. Rev. Lett.* **81**, 3395–3398 (1998).
25. Rudolph, M., Shinbrot, T. & Lueptow, R. M. A model of mixing and transport in wavy Taylor-Couette flow. *Physica D* **121**, 163–174 (1998).
26. Shinbrot, M. Fixed-point theorems. *Sci. Am.* **214**, 105–110 (1966).
27. Scott, D. R. Seismicity and stress rotation in a granular model of the brittle crust. *Nature* **381**, 592–595 (1996).
28. Pöschel, T. & Buchholtz, V. Complex flow of granular material in a rotating cylinder. *Chaos Solitons Fractals* **5**, 1901–1912 (1995).
29. Wolf, A., Swift, J. B., Swinney, H. L. & Vastano, J. A. Determining Lyapunov exponents from a time series. *Physica D* **16**, 285–317 (1985).

Acknowledgements. This work was supported by the International Fine Powder Research Institute, the NSF, the New Jersey Commission on Science and Technology, and Pfizer Pharmaceuticals. We thank A. Abad for technical assistance.

Correspondence and requests for materials should be addressed to F.J.M. (e-mail: muzzio@sol.rutgers.edu).

Self-organized growth of alloy superlattices

P. Venezuela*, J. Tersoff†, J. A. Floro‡, E. Chason‡§, D. M. Follstaedt‡, Feng Liu* & M. G. Lagally*

* Department of Materials Science and Engineering, University of Wisconsin, Madison, Wisconsin 53706, USA

† IBM Research Division, T. J. Watson Research Center, PO Box 218, Yorktown Heights, New York 10598, USA

‡ Sandia National Laboratories, Albuquerque, New Mexico 87185-1415, USA

Patterning in nature typically occurs through self-organization, and interest has developed recently in the use of such spontaneous processes to fabricate periodically structured materials at the nanometre scale. For example, ordered arrays of semiconductor ‘quantum dot’ particles (superlattices) have been created by deposition from a suspension¹, or by self-organization of diffusing atoms on surfaces² or in sequentially grown stacked layers³. The spontaneous formation of layered structures in epitaxial growth has also been reported, and attributed to the process of spinodal decomposition^{4,5}. Yet highly ordered layered superlattices, developed for applications in optoelectronics (and in future perhaps for thermoelectrics⁶), are created ‘by hand’ through the sequential deposition of two different materials. Here we show that superlattices can appear spontaneously during crystal growth of an alloy, as a consequence of the

§ Present address: Brown University, Division of Engineering, Providence, Rhode Island 02912, USA.

distribution of strain at surface step sites. When a strained alloy grows by 'step flow', the surface steps form periodic bunches⁷. We find that the resulting modulated strain field biases the incorporation of the respective alloy components at different steps in the bunch, leading to segregation and superlattice formation. We also present experimental observations (X-ray diffraction and electron microscopy) of a silicon-germanium alloy grown on silicon, which show clear evidence for the formation of such a self-organized structure.

Semiconductor devices are typically grown on a 'vicinal' surface, a staircase of atomically flat terraces separated by atomic-height steps. Atoms are deposited on the surface (by evaporation or vapour decomposition), and they diffuse as 'adatoms' on the terraces. Step motion arises from attachment or detachment of these adatoms at the step.

The adatom density on a terrace obeys the diffusion equation, subject to boundary conditions of equilibrium at the steps. Solving this equation gives the current of adatoms; the discontinuity in this current at a step gives the net attachment or detachment rate, and hence the step velocity. The physics of strain and alloying enters solely through the composition-dependent chemical potential at the steps, which determines the local adatom density and hence the adatom diffusion, step motion, and crystal growth⁸⁻¹⁰.

Recently Liu *et al.*⁷ showed that such growth leads to periodic patterns of step bunches on the crystal surface. Here we consider growth of an alloy, in which the elements differ in their size and surface mobility. The step bunches affect the surface elastic field, giving different strains at different steps. The smaller adatoms will be incorporated preferentially at steps having relatively compressive strain, and larger adatoms at relatively tensile steps^{9,11}. In addition, steps having lowest absolute strain are favoured overall, and atoms of the more mobile species can more readily reach these preferred steps¹⁰.

The equations describing alloy decomposition are nonlinear, and so inherently difficult to solve. We therefore consider the limit of only a small difference in size and/or mobility between the alloy elements, so that we can linearize the problem. In this case, and making use of prior results^{9,10}, we can write the evolution of surface morphology as

$$\dot{u}_m = \frac{1}{2}(u_{m+1} - u_{m-1}) + \bar{\lambda}^3 \Phi_m \quad (1)$$

Here $u_m = (x_m/L_{av}) - m - Ft$ is the step displacement, written for convenience in dimensionless units, relative to an ideal train of equispaced steps; x_m is the actual step position; L_{av} is the average step separation (fixed by the surface 'vicinal angle'), and F the total incident flux (the growth rate) in monolayers per unit time. (We do not include any step-edge diffusion barriers.) The corresponding dimensionless step velocity is $\dot{u}_m = du_m/d\tau$, where $\tau = Ft$ is the dimensionless time. For an alloy of two components, indexed by the subscript ν , $\lambda_\nu = [\exp((\bar{\mu}_\nu - E_\nu)/kT)D_\nu M \bar{\epsilon}^2 \beta h / 4FkT]^{1/3} / L_{av}$ characterizes the strain-driven diffusion for component ν . Here $\bar{\mu}_\nu$ is the average chemical potential, E_ν and D_ν are the energy and mobility of an adatom of component ν , T is the temperature, M is the product of the atomic volume and an elastic constant, $\bar{\epsilon}$ is the misfit strain, β is a ratio between elastic constants, and h is the step height. Only the average $\bar{\lambda} = (\lambda_\nu + \lambda_{-\nu})/2$ enters Equation (1), where the subscript $-\nu$ denotes the component other than ν . Finally,

$$\Phi_m = \frac{S_{m+1} - S_m}{u_{m+1} - u_m + 1} - \frac{S_m - S_{m-1}}{u_m - u_{m-1} + 1}$$

where $S_m = \sum_{n \neq m} (u_m + m - u_n - n)^{-1} - l_0^2 \sum_{n \neq m} (u_m + m - u_n - n)^{-3}$, and l_0 is the minimum-energy distance between two isolated steps, in units of L_{av} .

The dynamical evolution of the surface morphology is determined by numerical integration of the step velocity, equation (1). The behaviour is controlled by two parameters, l_0 and $\bar{\lambda}$. These

reflect respectively the short-range step repulsion and the long-range strain field (because the step changes the elastic boundary condition for the misfit strain).

The morphological evolution has been treated previously⁷. The new result which makes our simulation possible is the alloy decomposition at the steps:

$$\Delta c_\nu = \bar{c}_\nu \bar{c}_{-\nu} \alpha_\nu \Psi [I - \gamma \bar{\lambda}^3 A]^{-1} \quad (2)$$

Here Δc_ν gives the deviation from the average composition $\bar{c}_\nu$. Its m th component refers to step m . The m th component of Ψ is $\bar{\lambda}^3 \Phi_m / (1 + \dot{u}_m)$ for equal-sized atoms; I is the identity matrix; and A is a cyclic tridiagonal matrix with $A_{m,m \pm 1} = \pm (1 + \dot{u}_m)^{-1} (u_{m \pm 1} - u_m \pm 1)^{-1}$ and $A_{m,m} = -(A_{m,m+1} + A_{m,m-1})$. A detailed derivation of these results will be given elsewhere.

Given the surface morphology, there are two important (and competing) terms which determine the alloy decomposition. The difference in size and mobility of the two species are reflected in the parameter

$$\alpha_\nu = 4 \frac{\Delta \epsilon_\nu}{\bar{\epsilon}} + 3 \frac{\Delta \lambda_\nu}{\bar{\lambda}} \quad (3)$$

where $\Delta \epsilon_\nu / \bar{\epsilon}$ is the fractional misfit difference, with $\Delta \epsilon_\nu = (a_{-\nu} - a_\nu) / a_0$, a_ν and a_0 being the lattice parameters of component ν and of the substrate, respectively; and $\Delta \lambda_\nu / \bar{\lambda}$ is the fractional difference in mobility, relative to the average, where $\Delta \lambda_\nu = \lambda_\nu - \bar{\lambda}$. We note that the effects of atomic size and mobility difference can add or partially cancel, depending on the relative signs¹⁰.

The other important parameter is

$$\gamma = \bar{c}_\nu \bar{c}_{-\nu} 4 \frac{L_{av}}{\beta h M \bar{\epsilon}^2} g''_\nu \quad (4)$$

This acts in the opposite way from α_ν , to suppress decomposition. Here g''_ν is the second derivative, with respect to composition, of the free energy of mixing of the alloy. As long as the alloy is stable against spinodal decomposition, $g''_\nu > 0$ and so $\gamma > 0$. The larger γ is, the

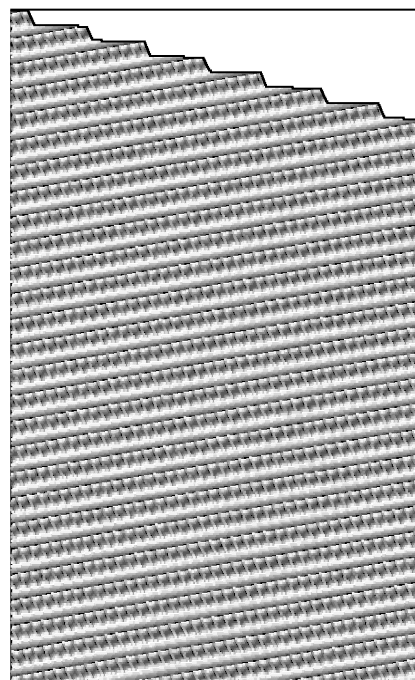


Figure 1 Cross-sectional view of simulated superlattice formation. The parameters used in this calculation are: $\bar{\lambda} = 0.446$, $\Delta \epsilon = 0$ and $l_0 = 0.192$. The vertical axis has been exaggerated for clarity. The black and white regions are enriched in one or the other component, while the grey regions have composition close to the average composition.

smaller the decomposition Δc_p . Other factors being equal, the alloy decomposition will be largest in systems where the critical temperature T_c for spinodal decomposition is not far below the growth temperature T , because $\gamma \rightarrow 0$ as $T \rightarrow T_c$.

An example of the calculated growth is shown in Fig. 1, for $l_0 = 0.192$, $\Delta\epsilon = 0$, and $\bar{\lambda} = 0.446$. The simulation included 120 steps, with periodic boundary conditions to eliminate end effects. The vertical axis in Fig. 1 has been exaggerated for clarity. The black and white regions are enriched in one or the other component, while the grey regions have composition close to the average value.

The result is an ordered pattern of composition modulation. We believe that this is the first actual calculation showing spontaneous superlattice formation in any system. Note the remarkable regularity of the structure, which appears to be independent of initial conditions and robust against noise. Increasing the growth rate reduces $\bar{\lambda}$, giving smaller bunches⁷ and hence a shorter superlattice period. It was speculated previously, on heuristic grounds, that such decomposition would occur at step bunches⁹. However, owing to the collective nature of the phenomenon, and the importance of mobility differences¹⁰, the actual results here differ dramatically from earlier speculations.

Because of the combination of short-range and long-range strains around each step, the total strain (and hence the chemical potential) is higher at the top and bottom of a step-bunch, and lower in the middle. The more mobile atoms are more successful at reaching the favoured mid-bunch sites, and so are preferentially incorporated

there. The atomic size similarly biases the incorporation of the two types of atoms¹⁰.

We note that in Fig. 1 the superlattice is misorientated in relation to the crystallographic plane (horizontal direction) by an angle which is comparable to the surface vicinal angle, but in the opposite direction. This is due to the complex step dynamics, in which free steps are continually ejected from the base of a bunch and then captured by the adjacent bunch⁷.

We have repeated the simulation for several values of γ . In general, the pattern of the decomposition reflects the surface morphology, and so is unaffected by γ . However, γ has a strong effect on the magnitude of the decomposition. Near the critical temperature T_c for spinodal decomposition, γ is small and the decomposition is quite large. But it is difficult to evaluate γ for specific real systems, because T_c is not accurately known.

In Fig. 2 we show measurements of a $\text{Si}_{0.84}\text{Ge}_{0.16}$ alloy grown by molecular beam epitaxy on Si(001) at 380 °C. The sample was grown without wafer rotation (periodic structures are sometimes created accidentally during growth by wafer rotation). The surface 'miscut', the deviation from (001) orientation, is too small to measure accurately, so the local miscut is dominated by the small deviations from planarity.

A cross-sectional transmission electron microscope (TEM) image is shown in Fig. 2a. The superlattice modulation, with a period of ~ 3 nm, is clearly visible despite the 'noisiness' of the image. (Viewing the image obliquely from the side enhances the clarity.)

X-ray diffraction provides a more quantitative measure of composition modulation (Fig. 2b). Comparison with calculated diffraction intensities suggests that the absolute composition modulation is of the order of 10%. Each peak in the diffracted intensity indicates the presence of some composition modulation with the period indicated. The multiple peaks indicate the presence of several distinct periods. All are around 3 nm, consistent with the TEM. A Fourier analysis of the TEM image (not shown here) also shows multiple periods separated by ~ 0.3 nm. We infer that the sample has microdomains, so that roughly speaking, each peak originates in a region having that periodicity.

We note that the separation between the principal diffraction peaks in Fig. 2b corresponds to an integer number of atomic double-layers (~ 0.27 nm). This confirms that the structure is truly self-organized, because an external perturbation would not give multiple periodicities differing by integer thicknesses.

It is not possible to prove definitely that the experiment reflects the same mechanism proposed in the theory. But because no other mechanism is known which could lead to such behaviour, the presumption is reasonable. Both our theory and our experiments are for systems that are stable against spinodal decomposition, and so are unrelated to previous observations^{4,5} or previous heuristic theoretical discussions^{12,13} for unstable systems. Moreover, the presence of multiple periodicities is consistent with the proposed mechanism. The average bunch size may vary across the sample owing to small differences in local surface miscut, flux or temperature. But bunches can differ only by integer numbers of steps; and if bunches of different size coexist they tend to "phase separate" into separate regions of uniform bunch size⁷. Such regions will have periods differing by integer numbers of atomic layers. (SiGe is well known to show step 'pairing', and the dominant spacing in Fig. 2 corresponds to the bilayer step height.)

There is still much to be understood in this system. We have not yet identified the precise experimental conditions for reproducibly growing superlattices. Also the quantitative predictions of the theory will be affected if we include other effects such as diffusion barriers, step permeability, and other step-repulsion mechanisms. But none of those would eliminate the basic morphological self-organization, and the resulting periodic modulation of the alloy composition. These appear to be very robust effects, as long as growth takes place in the step-flow mode described. □

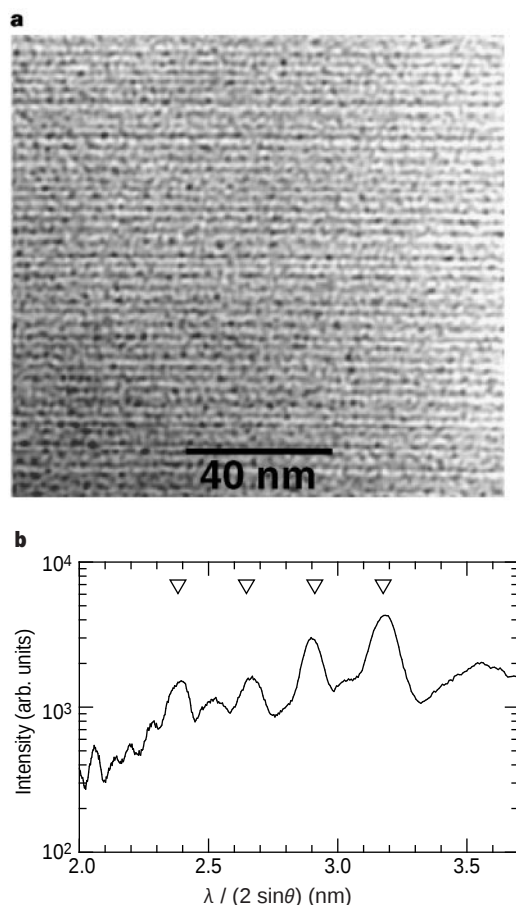


Figure 2 Observations of superlattice formation for a $\text{Si}_{0.84}\text{Ge}_{0.16}$ alloy grown on Si(001). **a**, Bright-field TEM image, [110] cross-section, obtained with (002) two-beam diffraction conditions. The top edge is towards the growth surface. Despite 'noise', horizontal layers with spacings around 3 nm are clearly visible. **b**, X-ray diffraction intensity (from θ - 2θ scan), versus periodicity (from X-ray wavelength λ and scattering angle θ). The arrow heads indicate successive major peaks differing by the bilayer step height.

Received 4 September; accepted 9 December 1998.

1. Murray, C. B., Kagan, C. R. & Bawendi, M. G. Self-organization of CdSe nanocrystallites into three-dimensional quantum dot superlattices. *Science* **270**, 1335–1338 (1995).
2. Shchukin, V. A., Ledentsov, N. N., Kop'ev, P. S. & Bimberg, D. Spontaneous ordering of arrays of coherent strained islands. *Phys. Rev. Lett.* **75**, 2968 (1995).
3. Tersoff, J., Teichert, C. & Lagally, M. G. Self-organization in growth of quantum dot superlattices. *Phys. Rev. Lett.* **76**, 1675–1678 (1996).
4. Petroff, P. M., Cho, A. Y., Reinhart, F. K., Gossard, A. C. & Wiegmann, W. Alloy clustering in Ga_{1-x}Al_xAs compound semiconductors grown by molecular beam epitaxy. *Phys. Rev. Lett.* **48**, 170–173 (1982).
5. Norman, A. G., Seong, T. Y., Ferguson, I. T., Booker, G. R. & Joyce, B. A. Structural studies of natural superlattices in group III–V alloy epitaxial layers. *Semicond. Sci. Technol.* **8**, S9–S15 (1993).
6. Mahan, G. D. & Woods, L. M. Multilayer thermionic refrigeration. *Phys. Rev. Lett.* **80**, 4016–4019 (1998).
7. Liu, F., Tersoff, J. & Lagally, M. G. Self-organization of steps in growth of strained films on vicinal substrates. *Phys. Rev. Lett.* **80**, 1268–1271 (1998).
8. Tersoff, J., Phang, Y. H., Zhang, Z. & Lagally, M. G. Self-bunching instability of vicinal substrates under stress. *Phys. Rev. Lett.* **75**, 2730–2733 (1995).
9. Tersoff, J. Stress-driven alloy decomposition during step-flow growth. *Phys. Rev. Lett.* **77**, 2017–2020 (1996).
10. Venezuela, P. & Tersoff, J. Alloy decomposition during growth due to mobility differences. *Phys. Rev. B* **58**, 10871–10874 (1998).
11. Guyer, J. E. & Voorhees, P. W. Morphological stability of alloy thin films. *Phys. Rev. Lett.* **74**, 4031–4034 (1995).
12. Barabási, A. Self-organized superlattices formation in II–IV and III–V semiconductors. *Appl. Phys. Lett.* **70**, 764–766 (1997).
13. Tersoff, J. Spinodal decomposition during step-flow growth. *Phys. Rev. B* **56**, R4394–R4397 (1997).

Acknowledgements. P.V. was supported by the Brazilian agency FAPESP.

Correspondence and requests for materials should be addressed to J.T. (e-mail: tersoff@us.ibm.com).

Non-aqueous supramolecular assembly of mesostructured metal germanium sulphides from (Ge₄S₁₀)⁴⁻ clusters

Mark J. MacLachlan, Neil Coombs & Geoffrey A. Ozin

Materials Chemistry Research Group, Department of Chemistry,
80 St George Street, Toronto, Canada M5S 3H6

Microporous materials have found extensive application as catalysts, ion-exchange media and sorbents^{1,2}. The discovery of mesoporous silica³ has opened the path to selective catalysis and separation of large molecules and to the synthesis of inorganic–organic composite materials, polymer mesofibres and semiconducting quantum dots^{4–7}. Various oxide-based mesoporous materials, such as TiO₂, ZrO₂, SnO₂, Al₂O₃, Nb₂O₅ and GeO₂, have been reported^{8–13}. A challenge for materials research is now to expand the scope of mesoporous materials beyond the oxides. Only a few non-oxide mesostructured composites, such as CdS, SnS₂ and CdSe, have been reported; they are usually synthesized by *ad hoc* hydrothermal methods or from aqueous solutions containing ill-defined species, and are often not well characterized^{14–16}. Here we report the rational synthesis of a new family of metal germanium sulphide mesostructured materials prepared by a non-aqueous surfactant-templated assembly of adamantanoid [Ge₄S₁₀]⁴⁻ cluster precursors. In the presence of quaternary alkylammonium surfactants, [Ge₄S₁₀]⁴⁻ anions in formamide solution self-organize with metal cations (Co²⁺, Ni²⁺, Cu⁺ and Zn²⁺) to create well ordered hexagonal metal germanium sulphide mesostructures, some having fibre-like morphologies with channels running down the long axis of the fibre. Materials of this genre could prove effective in applications as diverse as detoxification of heavy metals in polluted water streams, sensing of sulphurous vapours, and the formation of semiconductor quantum ‘anti-dot’ devices.

Open-framework metal germanium sulphides with ordered micropores are a class of materials in which the metal, chalcogenide and template molecules may be varied to ‘tune’ the pore size and properties of the resulting structure^{17–23}. In water, metal ions

assemble with anionic [Ge₄S₁₀]⁴⁻ adamantanoid clusters in the presence of tetramethylammonium (TMA) cations to give crystalline metal germanium sulphide microstructured materials¹⁹. We identified this approach as a modular assembly route to metal germanium sulphide mesostructured materials. But to achieve this goal the surfactant, [Ge₄S₁₀]⁴⁻ and metal precursors must be soluble and assemble into mesophases in water. Unfortunately, precursors like CTA₄Ge₄S₁₀ (where CTA is cetyltrimethylammonium) are sparingly soluble in water. With these restrictions in mind, we chose formamide as a solvent for the preparations because it will dissolve all of the reagents, and because the binary phase diagram of surfactants in formamide resembles that of water²⁴.

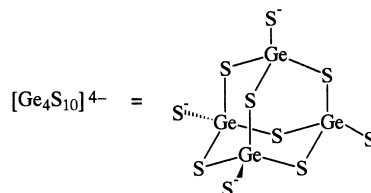


Figure 1a shows a typical powder X-ray diffractogram of the mesostructured Ni/[Ge₄S₁₀]⁴⁻ material prepared according to the procedures outlined in the Methods section. The product displays four reflections consistent with diffraction from a hexagonal unit cell with dimension $a = 39.5 \text{ \AA}$, attesting to the high degree of order of this material. Moreover, the diffraction pattern reveals no evidence for crystalline materials that would be expected from residual surfactant or dense phases. Powder X-ray diffractograms of the analogous Zn, Cu and Co materials are shown in Fig. 1b. Although they do not appear to be as well ordered as the Ni/[Ge₄S₁₀]⁴⁻ products, we attribute the broadening of the peaks and the overlap of the 110 and 200 reflections in these materials to small particle sizes (less than $\sim 200 \text{ nm}$). Indeed, transmission electron microscopy (TEM) suggests that many of these samples are better ordered than their nickel analogue, though the particles are clearly smaller.

TEM images of whole-mounted samples are shown in Fig. 1c–f. Hexagonal mesoscopic ordering of the materials extends over entire particles constituting the samples; overall, the samples appear homogeneous when viewed under the electron microscope. Energy-dispersive X-ray analysis of single mesostructured particles confirmed the presence of germanium, sulphur and the metal. Worm-like morphologies (Fig. 1e–f) have been observed by TEM.

Raman spectroscopy of the mesostructured materials confirms the presence of metal-linked adamantanoid clusters. Figure 2A, trace a, shows the Raman scattering from a sample of (CTA)₄Ge₄S₁₀ containing the molecular cluster²⁵. In particular, the spectrum displays two sets of modes due to Ge–S stretching: (1) terminal Ge–S_t modes between 340 and 480 cm⁻¹; and bridging Ge–S_b modes between 160 and 230 cm⁻¹. The symmetric Ge–S_t stretching mode observed at 342.0(±0.1) cm⁻¹ has a full-width at half-maximum (FWHM) of 7.6(±0.6) cm⁻¹. For comparison, the Raman spectrum of the crystalline TMA₂ZnGe₄S₁₀ microstructure, in which the [Ge₄S₁₀]⁴⁻ clusters are intact and linked by tetrahedral metal sites, is shown in Fig. 2A, trace b. In the crystalline framework structure, the symmetric mode shifts to 350.7(±0.2) cm⁻¹ but the linewidth of this peak remains virtually unchanged (FWHM, 8.8(±0.5) cm⁻¹). We note that weak modes attributed to Zn–S stretching are present between 250 and 320 cm⁻¹. Figure 2A, trace c, shows the Raman spectrum of a sample of the mesostructured Zn/[Ge₄S₁₀]⁴⁻ material. The symmetric Ge–S_t stretching mode is observed at 350.7(±0.2) cm⁻¹ with a FWHM of 19(±2) cm⁻¹. Moreover, the other Ge–S_t and Ge–S_b modes in the mesostructured material resemble those of the linked microstructure (Fig. 2A, trace b) but are much broader. A congested region of overlapping vibrational modes between 240 and 320 cm⁻¹ are Zn–S stretching modes, as observed in linked microstructured Zn/[Ge₄S₁₀]⁴⁻ materials. These results are consistent with metal-linked [Ge₄S₁₀]⁴⁻ clusters present in a number of sites, where the

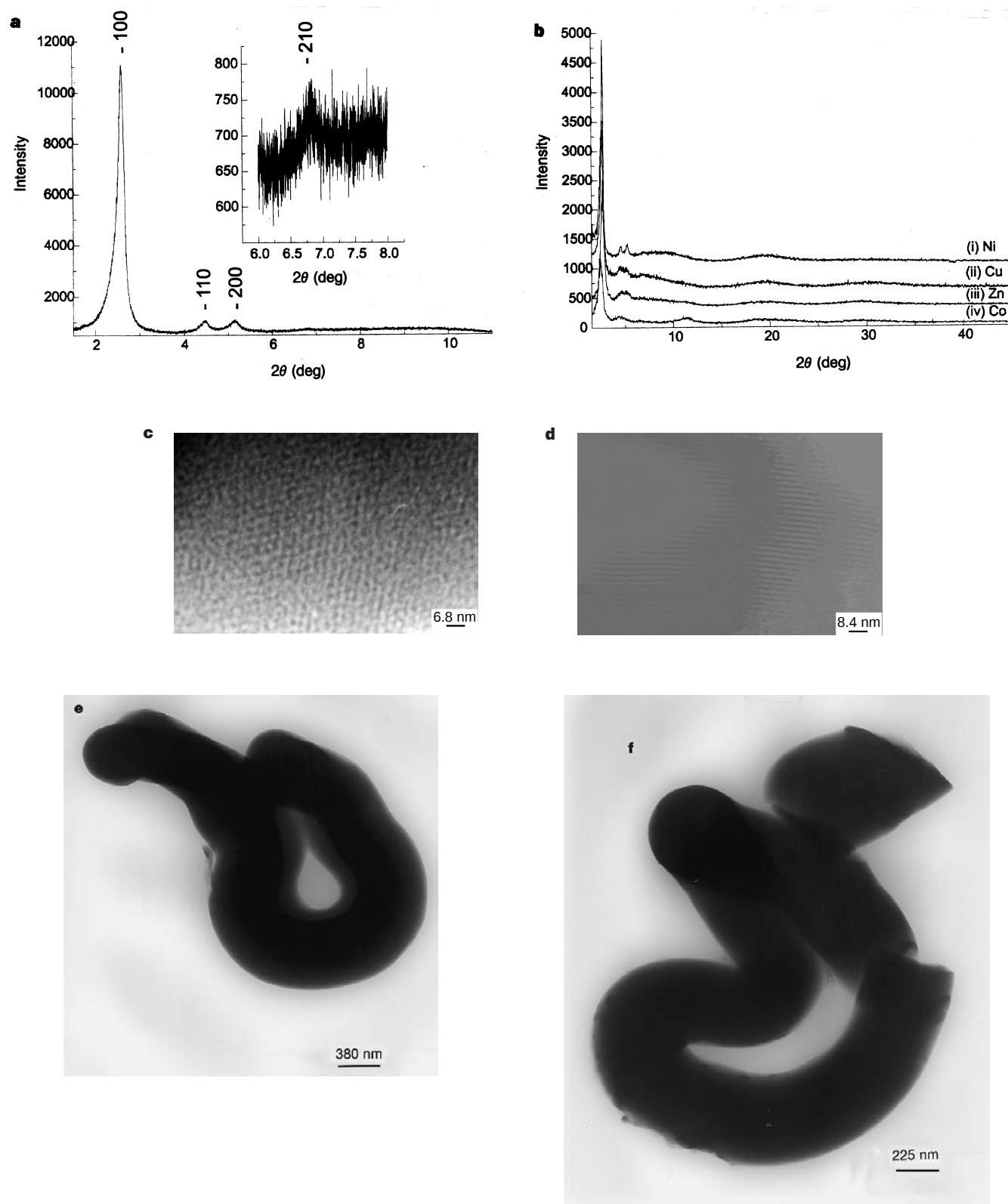


Figure 1 X-ray diffraction and TEM results. **a**, Powder X-ray diffraction (Siemens D5000, Cu-K α) of mesostructured Ni/[Ge₄S₁₀]⁴⁻ shows four reflections indexed to a hexagonal symmetry mesostructure (0.002° step size, 3.0-s step time). The inset depicts a scale expansion of the 6–8° 2 θ range showing the 210 reflection. Reflections (hkl) and their corresponding experimental and calculated d -spacings (Å) for a hexagonal unit cell ($a = 39.465$ Å) are as follows: (200), 34.18(1), 34.177; (110), 19.75(8), 19.732; (200), 17.19(8), 17.089; (210), 12.99(50), 12.918. **b**, Powder X-ray diffraction of mesostructured M/[Ge₄S₁₀]⁴⁻ materials (where M = Ni, Cu, Zn and

Co) (0.020° step size, 1.2-s step time). **c**, **d**, TEM images (Phillips 430, 100 kV) of mesostructured M/[Ge₄S₁₀]⁴⁻ show the long-range order present in the samples. **e**, **f**, TEM images of the Ni/[Ge₄S₁₀]⁴⁻ materials show the fibre morphologies observed in the samples. The worm-like structure (**e**) shows faceting that is expected for a hexagonal mesostructure. In addition, magnified TEM images confirmed that the hexagonal channels are aligned parallel to the longitudinal axis of the worm-like structures. (TEM images are of whole-mounted samples; embedding media used for sectioning the samples reacted with the materials.)

increased Raman line widths probably arise from variation in the coordination environment of the metal. By comparison with CTABr, the Raman spectrum of the mesostructured material shows the presence of CTA surfactant in the structure with little evidence for SH groups (Fig. 2B).

Solid-state magic-angle-spinning (MAS) ^1H and cross-polarization (CP) MAS ^{13}C NMR confirmed the presence of the CTA surfactant inside the materials. Figure 3a shows the ^{13}C NMR spectra of samples of $\text{Zn}/[\text{Ge}_4\text{S}_{10}]^{4-}$ and $\text{Cu}/[\text{Ge}_4\text{S}_{10}]^{4-}$ mesostructures. The spectra are virtually identical, showing the resonance due to the methyl groups on nitrogen (54 p.p.m.), the methylene attached to nitrogen (68 p.p.m.), and other resonances all consistent with the aliphatic hydrocarbon tail of the surfactant molecule. Furthermore, no resonance attributed to formamide ($\delta(^{13}\text{C}) = 167$ p.p.m.) is observed in the products.

Thermogravimetric analysis (TGA) also confirms the presence of surfactant in the materials. Figure 3b shows a typical TGA trace. A weight loss of ~ 40 – 50% is observed between room temperature and 500°C , with the onset of the main thermal event at ~ 200 – 215°C . Differential thermal analysis (DTA) shows two thermal events—a large weight loss at 250°C and a smaller one at 330°C . Both events are due to loss of CTA template from the materials, as observed in mesoporous silica. Indeed, pyrolysis mass spectrometry confirmed the existence of surfactant fragments (trimethylamine, cetyldimethylamine and 1-hexadecene) during decomposition. No mercaptans (for example, HSCH_3) or dialkylsulphides (such as

$\text{S}(\text{CH}_3)_2$) were observed by using this technique; this is in contrast to the crystalline $\text{TMA}_2\text{MGe}_4\text{S}_{10}$ microstructured frameworks, in which reaction of the TMA template with the metal sulphide induces elimination of $\text{S}(\text{CH}_3)_2$, causing framework collapse. This implies that framework collapse of the metal germanium sulphide mesostructure occurs by a different pathway from the microstructure.

Ten samples were analysed for metal, Ge, S, C, N and H content. Most of the analyses for the Zn, Co and $\text{Ni}/[\text{Ge}_4\text{S}_{10}]^{4-}$ samples were nearly $(\text{CTA})_2\text{M}_2\text{Ge}_4\text{S}_{10}$ with ~ 3 – 10 wt% unaccounted mass. This Ge:S ratio is expected only if the integrity of the clusters remained intact. We note that elemental analysis of a sample of mesostructured $\text{Cu}/[\text{Ge}_4\text{S}_{10}]^{4-}$ material revealed an empirical formula of nearly $(\text{CTA})_2\text{Cu}_4\text{Ge}_4\text{S}_{10}$, even though the same stoichiometry of metal: $[\text{Ge}_4\text{S}_{10}]^{4-}$ cluster was used in the preparations. We believe that the mesostructure consists of $[\text{Ge}_4\text{S}_{10}]^{4-}$ clusters linked by $\text{Cu}(\text{I})$ – $\text{Cu}(\text{I})$ dimers, reminiscent of that observed in the $\text{TMA}_2\text{Cu}_2\text{Ge}_4\text{S}_{10}$ microstructure²⁶.

After washing with acetone for 4 hours, a $\text{Zn}/[\text{Ge}_4\text{S}_{10}]^{4-}$ mesostructured material showed a drastic change in its thermal properties. The only mass loss was present at 335°C , and a total weight loss of only 29% was observed up to 500°C (compared with 47% in the untreated sample). This indicates that a significant portion ($\sim 60\%$) of the cationic CTA surfactant present in the material is charge-balancing an anionic framework, whereas the remainder ($\sim 40\%$) is present as space-filler to reinforce the structure, but can be removed with solvents.

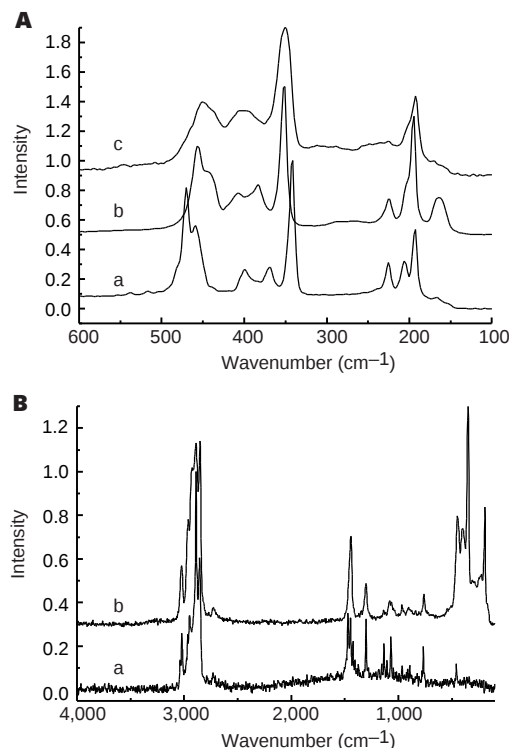


Figure 2 Raman spectra. **A**, $\text{CTA}_4\text{Ge}_4\text{S}_{10}$ (trace a), $\text{TMA}_2\text{ZnGe}_4\text{S}_{10}$ microporous crystalline framework (trace b), and mesostructured $\text{Zn}/[\text{Ge}_4\text{S}_{10}]^{4-}$ (trace c) show the Ge–S and Zn–S stretching modes in the 600 – 100 cm^{-1} region. **B**, Spectra of CTABr (trace a) and mesostructured $\text{Zn}/[\text{Ge}_4\text{S}_{10}]^{4-}$ (trace b) in the range $4,000$ – 100 cm^{-1} confirm the presence of CTA in the mesostructured material. We note that there are slight differences in the peak locations and linewidths between the pure surfactant and the surfactant inside the mesostructured $\text{Zn}/[\text{Ge}_4\text{S}_{10}]^{4-}$ material. Raman spectra were obtained on a Bomems MB-157 Fourier transform spectrometer with an InGaAs near-infrared laser (100 mW). Spectra (~ 100 scans) were collected from the samples contained in glass capillary tubes with a resolution of 4 cm^{-1} . The spectra fit well to Voigt amplitude lineshapes with quadratic backgrounds.

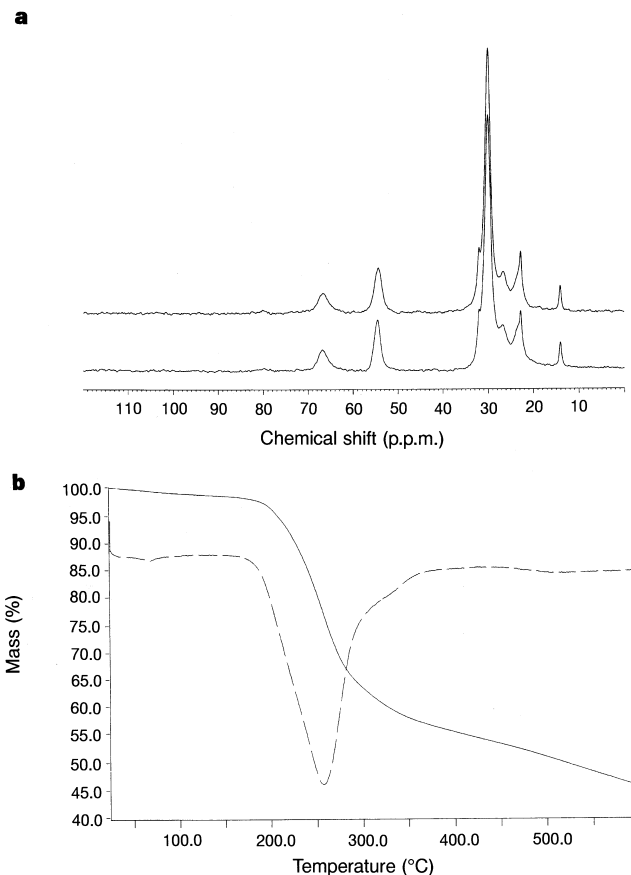


Figure 3 NMR and thermogravimetry results. **a**, Solid-state ^{13}C CP-MAS NMR spectra (Bruker DSX400, 100.62 MHz , 1.0 ms CP time, 10.0 s recycle delay, 4445 – 4625 scans) of mesostructured $\text{Cu}/[\text{Ge}_4\text{S}_{10}]^{4-}$ (top trace) and $\text{Zn}/[\text{Ge}_4\text{S}_{10}]^{4-}$ (bottom trace) confirm the presence of CTA in the product. We note that the spectra are virtually identical to, but distinct from, that of the pure surfactant. **b**, TGA (Perkin-Elmer TAC 7, 5°C min^{-1} , N_2) and DTA traces of mesostructured $\text{Ni}/[\text{Ge}_4\text{S}_{10}]^{4-}$, showing the thermal transition near 250°C .

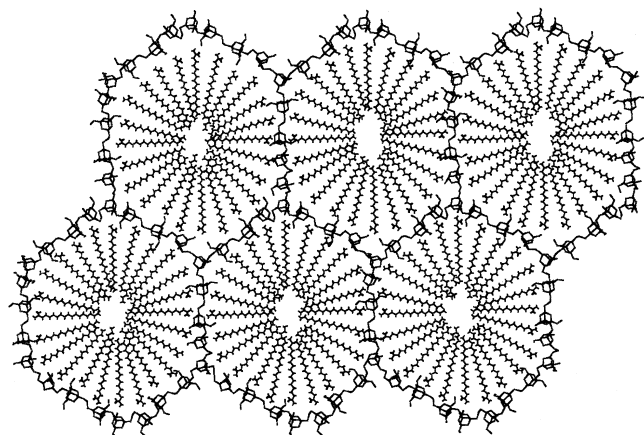


Figure 4 Proposed architecture of the mesostructured material. $\text{TMA}_4\text{Ge}_4\text{S}_{10}$ and CTABr are dissolved in formamide to give a viscous mesophase of the adamantanoid clusters charge-balanced by surfactant. On addition of a metal salt, the adamantanoid clusters link to form a supramolecular, hexagonal, mesostructured material in which the adamantanoid clusters are linked by metals in the structure. Although the mesostructure shown here is well defined, the microstructure remains speculative—we have no direct information yet about how the $[\text{Ge}_4\text{S}_{10}]^{4-}$ clusters are arrayed in the walls.

As the elemental analysis reveals nearly 2 CTA^+ surfactants per $[\text{Ge}_4\text{S}_{10}]^{4-}$ cluster in the framework, we believe that the initial micellar and liquid-crystalline phase may contain predominantly $[\text{Ge}_4\text{S}_{10}]^{4-}$ clusters which are charge-balanced by two CTA^+ cations in close proximity (compensated for by two TMA^+ cations). On addition of the transition-metal salt, the terminal sulphides of the adamantanoid clusters may displace ligands on the metal to form a mesostructured network (Fig. 4). Raman spectroscopy suggests that there are a variety of coordination environments for the transition metals; vacant coordination sites may be occupied by solvent or anions. In fact, when the blue $(\text{CTA})_2\text{Co}_2\text{Ge}_4\text{S}_{10}$ product is washed with water, it changes colour to green, probably due to exchanging water at the metal coordination sites. We speculate that the inorganic portion of the mesostructure is analogous to various micro- and mesostructured forms of SiO_2 , where tetrahedral SiO_2 is replaced by $\text{Ge}_4\text{S}_{10}\text{M}_{4/2}$. ($\text{M}_{4/2}$ denotes four two-connected S–M–S bridges between tetrahedral $[\text{Ge}_4\text{S}_{10}]^{4-}$ clusters.) This structural formula is consistent with the metal:germanium analysis in the materials.

A variety of transition metals can be used to form these mesostructured metal germanium sulphide materials. As germanium selenide and tin sulphide adamantanoid precursors are available, it might be possible to ‘tune’ the structures. The mesostructured metal germanium sulphides reported here, formed by supramolecular cluster assembly, have a highly flexible and adjustable composition which bodes well for a range of perceived applications, such as sulphur-tolerant catalysts for hydrocracking, hydrogenation and hydrodesulphurization, chemical sensing and removal of heavy metals from contaminated water. □

Methods

General procedure. A formamide solution (80 °C) of the transition-metal salt (CoSO_4 , $\text{Ni}(\text{NO}_3)_2$, ZnCl_2 , CuCl) is added slowly to a solution of $(\text{TMA})_4\text{Ge}_4\text{S}_{10}$ and a cationic surfactant (such as CTABr) in formamide (5–50 wt% surfactant), giving an immediate precipitate as the mesostructured material assembles. (The precipitation does not occur in the absence of transition metals, indicating that they are necessary to form the mesostructure.) Once the addition is complete, the mixture is shaken for a few seconds and aged at 80 °C for ~12–48 h. The powder is isolated on a Buchner funnel, washed with formamide (80 °C) then water (80 °C) to remove excess surfactant and reagents, and dried under ambient conditions before characterization. Through this route, gram quantities of the materials can be readily and reproducibly prepared.

A representative synthesis. A solution of 1.00 g $\text{TMA}_4\text{Ge}_4\text{S}_{10}$ and 1.50 g CTABr in 14 ml of formamide was formed at 120 °C and cooled to 80 °C. A solution of 642 mg $\text{Ni}(\text{NO}_3)_2 \cdot 6\text{H}_2\text{O}$ in 12 ml of formamide at 80 °C was added dropwise over 1 min to the solution of $\text{TMA}_4\text{Ge}_4\text{S}_{10}$ /CTABr, giving an immediate brown precipitate. The mixture was aged at 80 °C for 16 h, filtered, washed with copious amounts of formamide (80 °C) and water (80 °C), and dried under ambient conditions. Yield: 1.37 g of brown powder. A similar procedure was used for the other materials, substituting CuCl , ZnCl_2 and $\text{Co}(\text{SO}_4)_2$ for the $\text{Ni}(\text{NO}_3)_2$ used in the procedure. A microanalysis of a $\text{Zn}/[\text{Ge}_4\text{S}_{10}]^{4-}$ sample was: Zn, 9.62%; Ge, 20.2%; S, 22.2% ($\pm 5\%$); C, 35.3%; H, 6.8%; N, 2.2%; Cl, <0.01%; Br, <0.01%; giving an overall stoichiometry of $\text{Zn}_{2.11}\text{Ge}_4\text{S}_{9.95}\text{C}_{42.2}\text{H}_{97.0}\text{N}_{2.26}$, close to an ideal $(\text{CTA})_2\text{Zn}_2\text{Ge}_4\text{S}_{10}$. Karl–Fischer titration showed 4.95 wt% H_2O . An average of all of the analysed samples formed without sulphate (8 samples) gave a Ge:S ratio of 4:9.4 (± 0.8), which agrees well with the theoretical ratio of 4:10. Other surfactants, such as cetylpyridinium chloride and octadecyltrimethylammonium bromide, readily template the formation of the materials. When CuCl was used, the solution was filtered before addition to the $\text{TMA}_4\text{Ge}_4\text{S}_{10}$ /CTABr solution to remove undissolved salt.

Received 14 September; accepted 26 November 1998.

1. Szostak, R. *Handbook of Molecular Sieves* (Van Nostrand Reinhold, New York, 1992).
2. Breck, D. W. *Zeolite Molecular Sieves: Structure, Chemistry, and Use* (Wiley-Interscience, Toronto, 1974).
3. Kresge, C. T., Leonowicz, M., Roth, W. J., Vartuli, J. C. & Beck, J. C. Ordered mesoporous molecular sieves synthesized by a liquid-crystal template mechanism. *Nature* **359**, 710–712 (1992).
4. Sellinger, A. *et al.* Continuous self-assembly of organic–inorganic nanocomposite coatings that mimic nacre. *Nature* **394**, 256–260 (1998).
5. Wu, C. & Bein, T. Conducting polyaniline filaments in a mesoporous channel host. *Science* **264**, 1757–1759 (1994).
6. Maschmeyer, T., Rey, F., Sankar, G. & Thomas, J. M. Heterogeneous catalysts obtained by grafting metallocene complexes onto mesoporous silica. *Nature* **378**, 159–162 (1995).
7. Chomski, E., Dag, Ö., Kuperman, A., Coombs, N. & Ozin, G. A. New forms of luminescent silicon:silica composite mesostructures. *Chem. Vap. Deposition* **2**, 8–13 (1996).
8. Hue, Q. S. *et al.* Generalized synthesis of periodic surfactant/inorganic composite materials. *Nature* **368**, 317–321 (1994).
9. Antonelli, D. M. & Ying, J. Y. Synthesis of a stable hexagonally packed mesoporous niobium oxide molecular sieve through a novel ligand-assisted templating mechanism. *Angew. Chem. Int. Edn. Engl.* **35**, 426–430 (1996).
10. Bagshaw, S. A. & Pinnavaia, T. J. Mesoporous alumina molecular sieves. *Angew. Chem. Int. Edn. Engl.* **35**, 1102–1105 (1996).
11. Tian, Z.-R. *et al.* Manganese oxide mesoporous structures: mixed-valent semiconducting catalysts. *Science* **276**, 926–930 (1997).
12. Ciesla, U., Schacht, S., Stucky, G. D., Unger, K. K. & Scüth, F. Formation of a porous zirconium oxo phosphate with a high surface area by a surfactant-assisted synthesis. *Angew. Chem. Int. Edn. Engl.* **35**, 541–543 (1996).
13. Yang, P., Zhao, D., Margolese, D. I., Chmelka, B. F. & Stucky, G. D. Generalized syntheses of large-pore mesoporous metal oxides with semicrystalline frameworks. *Nature* **396**, 152–155 (1998).
14. Braun, P. V., Osenar, P. & Stupp, S. I. Semiconducting superlattices templated by molecular assemblies. *Nature* **380**, 325–328 (1996).
15. Fröba, M. & Oberender, N. First synthesis of mesostructured thiogermanates. *Chem. Commun.* 1729–1730 (1997).
16. Neeraj & Rao, C. N. R. Metal chalcogenide-organic nanostructured composites from self-assembled organic amine templates. *J. Mater. Chem.* **8**, 279–280 (1998).
17. Bedard, R. L., Vail, L. D., Wilson, S. T. & Flanigen, E. M. Hydrocarbon conversion process using crystalline microporous metal sulfide compositions. US Patent 4,933,068 (1990).
18. Bedard, R. L., Wilson, S. T., Vail, L. D., Bennett, J. M. & Flanigen, E. M. in *Zeolites: Facts, Figures, Future* (eds Jacobs, P. A. & van Santen, R. A.) 375–387 (Elsevier, Amsterdam, 1989).
19. Yaghi, O. M., Sun, Z., Richardson, D. A. & Groy, T. L. Directed transformation of molecules to solids: synthesis of a microporous sulfide from molecular germanium sulfide cages. *J. Am. Chem. Soc.* **116**, 807–808 (1994).
20. Tan, K., Darovsky, A. & Parise, J. B. Synthesis of a novel open-framework sulfide, $\text{CuGe}_2\text{S}_5 \cdot (\text{C}_2\text{H}_5)_4\text{N}$, and its structure solution using synchrotron imaging plate data. *J. Am. Chem. Soc.* **117**, 7039–7040 (1995).
21. Sheldrick, W. S. & Wachhold, M. Solventothermal synthesis of solid-state chalcogenidometalates. *Angew. Chem. Int. Edn. Engl.* **36**, 206–224 (1997).
22. Achak, O., Pivan, J. Y., Maunay, M., Louër, M. & Louër, D. Structure refinement by the Rietveld method of the thiogermanates $[(\text{CH}_3)_4\text{N}]_2\text{MGe}_4\text{S}_{10}$ (M = Fe, Cd). *J. Solid State Chem.* **121**, 473–478 (1996).
23. Bowes, C. L. *et al.* Thermally stable self-assembling open-frameworks: isostructural Cs^+ and $(\text{CH}_3)_4\text{N}^+$ iron germanium sulfides. *Chem. Ber.* **129**, 283–287 (1996).
24. Wörnheim, T. & Jönsson, A. Phase diagrams of alkyltrimethylammonium surfactants in some polar solvents. *J. Colloid Interface Sci.* **125**, 627–633 (1988).
25. Bonhomme, F. & Kanatzidis, M. G. Structurally characterized mesostructured hybrid surfactant-inorganic lamellar phases containing the adamantane $[\text{Ge}_4\text{S}_{10}]^{4-}$ anion: synthesis and properties. *Chem. Mater.* **10**, 1153–1159 (1998).
26. Bowes, C. L. *et al.* Dimetal linked open frameworks: $[(\text{CH}_3)_4\text{N}]_2(\text{Ag}_2, \text{Cu}_2)\text{Ge}_2\text{S}_{10}$. *Chem. Mater.* **8**, 2147–2152 (1996).

Acknowledgements. We thank R. L. Bedard for helpful discussions, critical reading of the manuscript and elemental analyses of the materials. We thank E. Chomski for creative and technical assistance with Fig. 4. G.A.O. acknowledges the Killam Foundation for the award of an Isaac Walton Killam research fellowship (1995–97). M.J.M. is indebted to NSERC for a post-graduate scholarship (1995–99). This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) and UOP.

Correspondence and requests for materials should be addressed to G.A.O. (e-mail: gozin@alchemy.chem.utoronto.ca).

124,000-year periodicity in terrestrial vegetation change during the late Pliocene epoch

K. J. Willis*, A. Kleczkowski† & S. J. Crowhurst*

* Godwin Laboratory, University of Cambridge, Free School Lane, Cambridge CB2 3RS, UK

† Department of Plant Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EA, UK

The late Pliocene (~3–2.6 million years ago) is an interval of exceptional interest for understanding the Earth's climate system. It was a time of progressive global cooling, resulting in the growth of large terrestrial ice sheets and the initiation of extensive Northern Hemisphere glaciation^{1,2}. The build up of the ice sheets was cyclical and apparently paced by the orbitally driven oscillations in incoming solar radiation (Milankovitch cycles) at periods of approximately 41 kyr (obliquity) and 23–19 kyr (precession). Here we present a high-resolution continental record of late Pliocene climate change, detailing the response of terrestrial vegetation to this interval of dramatic global environmental change. The annually laminated sequence of lake sediments from Pula maar, in Hungary, represents approximately 320 kyr of accumulation between ~3.0 and 2.6 million years ago. Spectral analyses of the record indicate terrestrial responses to incoming solar radiation at obliquity and precession periodicities, but the strongest response appears at a period of ~124 kyr. Calculations indicate that variations in insolation forcing at this periodicity

were negligible at this time. The Pula record thus demonstrates that internally driven nonlinear responses of the climate system, at a period of ~124 kyr, were at least as important as external forcing at the orbital frequencies of precession and obliquity in driving late Pliocene large-scale environmental change.

Pula maar is a basaltic tuff maar-type crater situated in central Hungary ~10 km north of Lake Balaton. When first formed, the crater consisted of steep basaltic tuff walls containing a lake estimated to have been up to 40 m deep⁴. The finely laminated structure of the sedimentary sequence suggests that the lake was current-free and non-agitated. K/Ar dating indicates that basin formation occurred at $\sim 3.82 \pm 0.93$ Myr (ref. 5) although it is probable that accumulation of laminated sediments into the basin did not occur until ~3.0 Myr; magnetic polarity reversal data (measured at 32-cm intervals throughout the core) have revealed that the sedimentary sequence is of normal polarity. Counts of the laminations indicate that the sequence covers ~320 kyr. The geomagnetic polarity timescale indicates that the closest interval of normal polarity to 3.82 Myr that lasted >320 kyr is located in the Gauss epoch (polarity chron C2An.1n), dated between 3.05 and 2.60 Myr (ref. 6). The top of the sequence, which extrapolates to a date of ~2.67 Myr, is capped with loess. This date corresponds to a time when widespread loess deposition occurred throughout the Northern Hemisphere⁷. The loess effectively sealed the basin, thus preserving a sedimentary sequence spanning 320 kyr in the late Pliocene.

The core was collected in continuous 1.5-m sections using wire-line drilling technology. Laminations were counted at set intervals throughout, using a computerized stage attached to a tree-ring counter, and subsamples (1 cm³) were collected for the analyses of fossil pollen at a regular sampling interval of 32 cm. 110 samples were analysed for fossil pollen; the results are plotted as a pollen

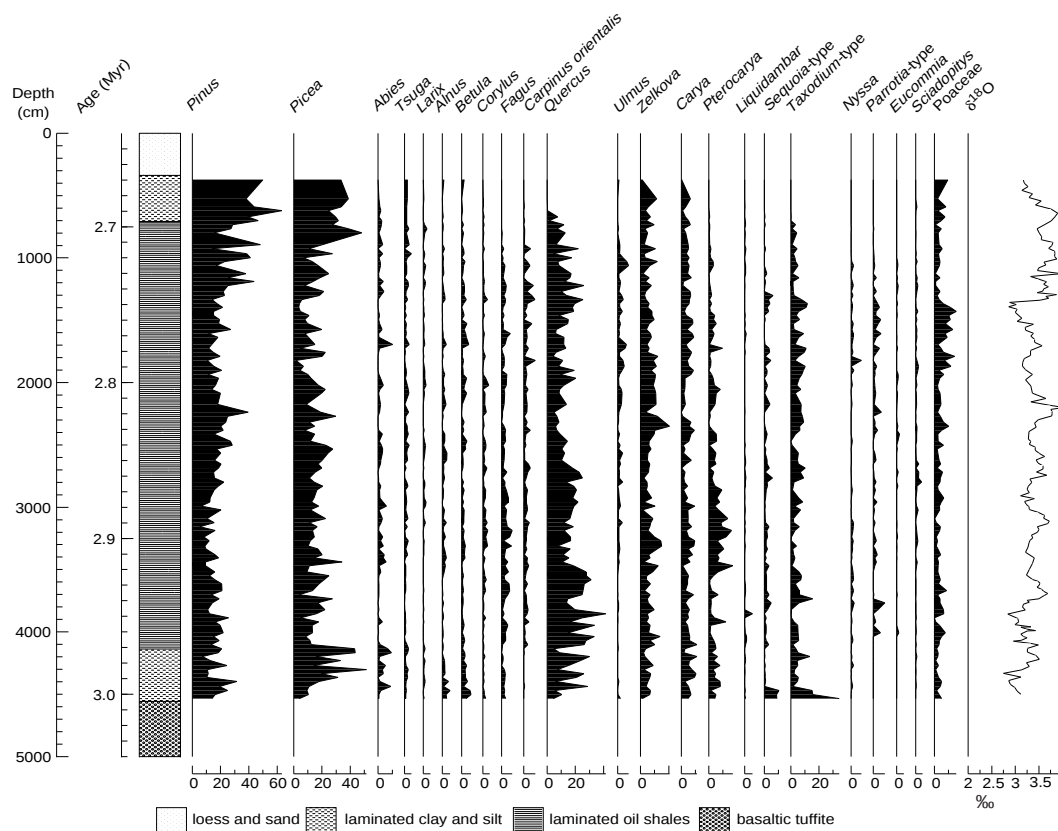


Figure 1 Pollen percentage diagram from Pula maar, Hungary. Data represent the percentages of total land pollen plotted against depth. Also included is an estimated age/depth scale based on K/Ar data, magnetic polarity reversal data

and counting of laminations. The pollen sampling interval represents approximately 1 sample every 2.5 kyr. Also shown are the benthic oxygen isotope data from ODP Site 659¹⁷ covering this time interval.

percentage diagram (Fig. 1). To determine the main direction of variance in the pollen percentage data set, the results were analysed using principal components analysis⁸ (Fig. 2).

The laminated layers, which are visible as light/dark layers in the sediment, account for over 80% of the sequence. The dark layers are predominantly composed of organic-rich clay with the algae *Botryococcus braunii*; the light layers are a carbonate precipitate. Previous work^{4,9,10} has demonstrated that the organic-rich clay layer represents winter accumulation, whereas the carbonate precipitates represent accumulation in spring and summer. One light layer plus one dark layer therefore represents a single year's accumulation. Counting the laminae indicates that throughout the core there is a remarkably regular sediment accumulation rate of ~80 yr (light/dark layers) per cm. The total core therefore represents ~320 kyr of annual accumulation between ~3.0 and 2.67 Myr ago, with a pollen sampling interval of approximately one sample every 2.5 kyr. Concentration and preservation of the pollen in this sequence is exceptional (~50,000 grains cm⁻³).

Pollen results (Fig. 1) indicate that during the late Pliocene there were significant fluctuations in the vegetation, both in terms of

population increases and decreases. Principal components analysis (Fig. 2) reveals that the primary direction of variance in the pollen percentage data is between those types found in the present-day northern boreal forest (for example, *Pinus*, *Tsuga*, *Larix*, *Alnus* and *Betula*) and those types found in sub-tropical/temperate forests (for example, *Sequoia*, *Nyssa*, *Eucommia*, *Pterocarya*, *Sciadopitys*, *Taxodium*, *Fagus*, *Quercus*, *Ulmus* and *Corylus*). The overall long-term trend indicates an increasing presence of boreal vegetation, culminating in a threshold-like transition between ~2.721 and 2.676 Myr ago when there was the disappearance of sub-tropical taxa (such as *Sequoia*, *Nyssa*, *Eucommia*, *Pterocarya* and *Sciadopitys*) (Fig. 3). Their disappearance represents the final demise of many of these species in northern Europe^{11–13}. The age of this transition coincides with the first extreme (glacial) excursion in the marine oxygen-isotope record, indicating a close link between vegetation change and the global development of large terrestrial ice sheets in the Northern Hemisphere².

Blackman–Tukey¹⁴ cross-spectral analysis was performed on boreal and sub-tropical pollen percentage values (Fig. 4) against calculated summer insolation for 47° N for the period 3.003–

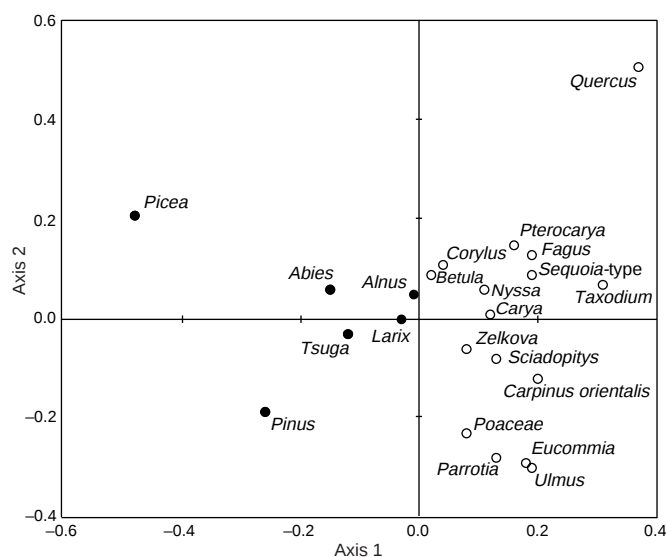


Figure 2 Principal components analysis (PCA) covariance biplot⁸ of pollen taxa from Pula maar. Only the first two axes were significant, accounting for ~60% of the total variance. The primary direction of variance in the Pula pollen data set is between those types found in the present-day northern boreal forest (for exam-

ple, *Pinus*, *Picea*, *Tsuga*, *Larix*, *Alnus* and *Betula*) and those types found in sub-tropical/temperate forests (such as *Sequoia*, *Nyssa*, *Eucommia*, *Pterocarya*, *Sciadopitys*, *Taxodium*, *Fagus*, *Quercus*, *Ulmus* and *Corylus*).

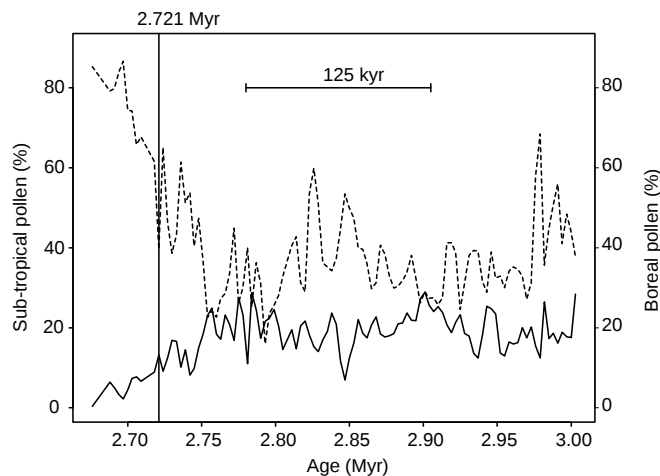


Figure 3 Sum of percentage pollen data from Pula Maar plotted against time. The data (sub-tropical (smooth line) and boreal (dashed line)) were interpolated using a 9-kyr gaussian window at 3-kyr intervals. The transition at 2.721 Myr represents

a non-stationary shift in the mean and is associated with the final demise of the sub-tropical assemblage.

2.721 Myr ago¹⁵. To reduce statistical biasing from dependences, the sub-tropical and boreal data sets were transformed before the spectral analysis using $\log(x + 1)$ (ref. 16).

Cross-spectral analysis of the boreal and sub-tropical percentage data (Fig. 4) indicates that in the Hungarian late Pliocene, vegetation was responding to climate change linked to orbital periodicities, in particular precession (19–23 kyr) and obliquity (41 kyr). The boreal and tropical pollen percentages are exactly out of phase with each other at these Milankovitch frequencies, suggesting that both taxonomic groups are responding with identical lags to the same forcing. It seems reasonable to assume that the sub-tropical taxa are increasing in response to increased Northern Hemisphere insolation while the boreal taxa are showing the opposite behaviour. Coherences are stronger with the 41- and 23–19-kyr periods in insolation than with equivalent periodicities in the benthic oxygen-isotope record¹⁷, suggesting that the flora are responding more

immediately to insolation than to global ice-volume changes at orbital timescales.

The boreal sequence is coherent at >80% confidence with orbital obliquity (41-kyr) and precession (23–19-kyr) components. Obliquity-related variations also dominate the marine oxygen-isotope sequence at this time interval, and it is reasonable to assume that the climate variations evident in the boreal assemblage are in some way associated with the same underlying mechanisms. Small ice sheets started to accumulate in the Northern Hemisphere ~10 Myr ago¹⁸, with progressive intensification from ~2.75 Myr ago^{1,2}. Superimposed on this long-term accumulation of terrestrial ice are clear 41-kyr ice-volume cycles paced by astronomically driven variations in insolation¹⁹. At least some of the variations seen in the boreal vegetation are therefore likely to be associated with variations in temperature associated with these ice-volume fluctuations.

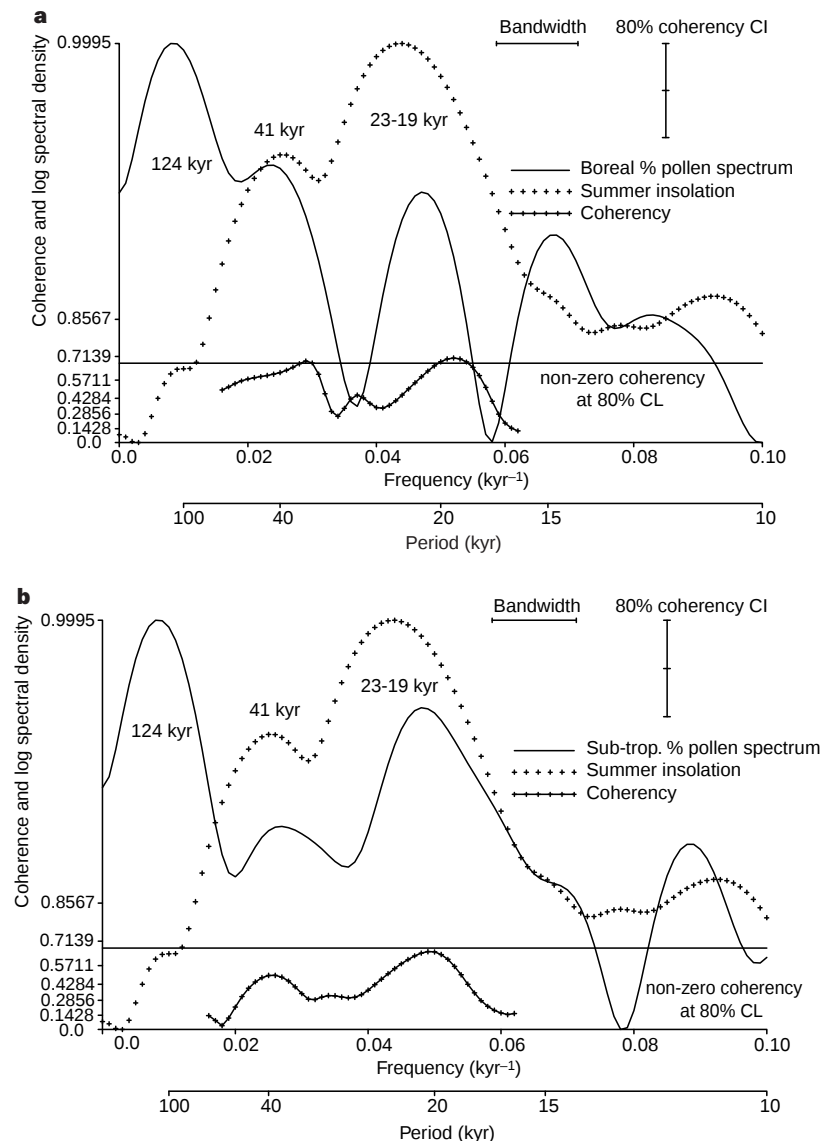


Figure 4 Cross-spectral comparison of boreal pollen percentage data with calculated summer insolation. Calculations are for July at 47° N, 2.721–3.003 Myr (ref. 15). Pollen types included in the analyses were *Pinus*, *Abies*, *Picea*, *Tsuga*, *Larix*, *Alnus* and *Betula*. The pollen percentage data were interpolated using a 9-kyr gaussian window at 3-kyr intervals and then transformed by $\log(x + 1)$. The linear trend was also removed. Spectral densities (variance/frequency) are normalized and plotted on log scales. The coherence spectrum (y-axis; solid line with crosses) is plotted on a hyperbolic arctangent scale. The solid horizontal

line denotes test-statistics for non-zero coherence at the 80 percent confidence level (CL). The spectral analysis was performed using the programme ARAND (P. J. Howell, Brown University). The two records are coherent both at the 19–23-kyr and the 41-kyr periodicities. **b**, As **a** but for sub-tropical pollen percentage data. Pollen types included in the analyses were *Eucommia*, *Pterocarya*, *Liquidambar*, *Sequoia*-type, *Nyssa*, *Carya*, *Elaeagnus*, *Parrotia*-type. The two records are coherent at the 19–23-kyr periodicity.

Spectral analysis of the sub-tropical percentage data (Fig. 4) indicates coherence with the 23–19-kyr oscillations (precession) just below the 80% confidence interval. The effects of the seasonality and precipitation produced by the increased amplitude of the seasonal cycle of solar radiation appear to have had a significant influence on the distribution and composition of the sub-tropical vegetation in the Hungarian late Pliocene.

The pollen spectrum of both taxonomic groups (that is, boreal and sub-tropical) is, however, dominated by a strong low-frequency component of ~124 kyr. Even though there are only at most three cycles to be resolved in this data set, such variance is clearly visible in the record (Fig. 3). This is of particular significance, as the 124-kyr peak is non-existent in the calculated insolation forcing (Fig. 4)¹⁵.

There is some evidence from oceanic records of significant environmental change at the 95–124-kyr interval. Dust data in deep sea cores exhibit a strong response at this period in the time interval considered here²⁰. Changes in atmospheric dust content are thought to reflect directly changes in continental aridity and terrestrial vegetation cover²¹, but until now there has been little direct pollen evidence to support this suggestion. The 124-kyr variance observed in our pollen record strongly supports the link between the dust content observed in marine cores and terrestrial vegetation change.

The results from Pula thus indicate that in addition to forcing at the orbital frequencies of precession and obliquity, internally driven nonlinear responses of the climate system at a period of ~124 kyr were as important (if not more important) in driving terrestrial vegetation dynamics and were presumably associated with this broad-scale environmental change. This terrestrial sequence provides the basis for beginning to understand the physical relationships between vegetation, ice volume and insolation forcing during a critical period in the Earth's climate system. □

Received 12 May; accepted 7 December 1998.

- Shackleton, N. J., Hall, M. A. & Pate, D. *Proc. ODP Sci. Res.* **138**, 337–345 (1995).
- Maslin, M. A., Haug, G. H., Sarntheim, M. & Tiedemann, R. The progressive intensification of northern hemisphere glaciation as seen from the North Pacific. *Geol. Rundsch.* **85**, 452–465 (1996).
- Zimmerle, W. in *Paleolimnology of European Maar Lakes* (eds Negendank, J. F. W. & Zolitschka, B.) 1–511 (Springer, Berlin, 1993).
- Ravasz, C. in *Annual Report of the Hungarian State Geological Institute for 1974* 221–245 (Hungarian Geol. Inst., Budapest, 1976).
- Balogh, K. et al. in *Annual Report of the Hungarian State Geological Institute for 1980* 243–260 (Hungarian Geol. Inst., Budapest, 1982).
- Cande, S. C. & Kent, D. V. A new geomagnetic polarity timescale for the Late Cretaceous and Cenozoic. *J. Geophys. Res.* **97**, 13917–13951 (1992).
- Kukla, G. Loess stratigraphy in central China. *Quat. Sci. Rev.* **6**, 191–219 (1987).
- Birks, H. J. B. & Gordon, A. D. *Numerical Methods in Quaternary Pollen Analysis* (Academic, London, 1985).
- Hajós, M. in *Annual Report of the Hungarian State Geological Institute for 1988* 5–13 (Hungarian Geol. Inst., Budapest, 1989).
- Heinz, T., Rein, B. & Negendank, J. F. W. in *Paleolimnology of European Maar Lakes* (eds Negendank, J. F. W. & Zolitschka, B.) 149–161 (Springer, Heidelberg, 1993).
- van der Hammen, T., Wijmstra, T. A. & Zagwijn, W. H. in *The Late Cenozoic Glacial Ages* (ed. Turekian, K. K.) 391–424 (Yale Univ. Press, New Haven, 1971).
- Tallis, J. H. *Plant Community History. Long-term Changes in Plant Distribution and Diversity* (Chapman & Hall, London, 1991).
- Suc, J. P. et al. Zanclean (Brunsumian) to early Piacenzian (early-middle Reuverian) climate from 4° to 54° north latitude (West Africa, West Europe and West Mediterranean areas). *Meded. Rijks Geol. Dienst* **52**, 43–56 (1995).
- Jenkins, G. M. & Watts, D. G. *Spectral Analysis and Its Applications* (Holden Day, San Francisco, 1968).
- Laskar, J., Joutel, F. & Boudin, F. Orbital, precessional and insolation quantities for the Earth from 20 Myr to 10 Myr. *Astron. Astrophys.* **270**, 522–533 (1993).
- Aitchinson, J. & Brown, J. A. C. *The Lognormal Distribution* (Cambridge Univ. Press, 1969).
- Tiedemann, R., Sarntheim, M. & Shackleton, N. J. Astronomic timescale for the Pliocene Atlantic $\delta^{18}\text{O}$ and dust records of Ocean Drilling Program Site 659. *Paleoceanography* **9**, 619–638 (1994).
- Thiede, J. & Myhre, A. M. *Proc. ODP Sci. Res.* (eds Thiede, J., Myhre, A. M., Firth, J. V., Johnson, G. L. & Ruddiman, W. F.) 645–659 (1996).
- Cronin, T. M., Raymo, M. E. & Kyle, K. P. Pliocene (3.2–2.4 Ma) ostracode faunal cycles and deep ocean circulation, North Atlantic Ocean. *Geology* **24**, 695–698 (1996).
- Tiedemann, R., Sarntheim, M. & Shackleton, N. J. Astronomic timescale for the Pliocene Atlantic $\delta^{18}\text{O}$ and dust flux records of Ocean Drilling Program site 659. *Paleoceanography* **9**, 619–638 (1994).
- Rea, D. K. The paleoclimatic record provided by eolian deposition in the deep sea: the geologic history of wind. *Rev. Geophys.* **32**, 159–195 (1994).

Acknowledgements. We thank K. D. Bennett, S. Clemens, M. Chapman, C. A. G. Gilligan, C. Heusser, M. F. Loutre, J. C. Ritchie, N. J. S. Shackleton and L. P. Zhou for comments and discussions on this manuscript, and B. Goddard for help with the figures. This work was funded by a Royal Society University Research Fellowship (K.W.) and a King's College Senior Research Fellowship (A.K.).

Correspondence and requests for materials should be addressed to K.J.W. (e-mail: kathy.willis@geog.ox.ac.uk).

Relative impacts of human-induced climate change and natural climate variability

Mike Hulme*, Elaine M. Barrow*, Nigel W. Arnell†, Paula A. Harrison‡, Timothy C. Johns§ & Thomas E. Downing‡

* Climatic Research Unit, School of Environmental Sciences, University of East Anglia, Norwich NR4 7TJ, UK

† Department of Geography, University of Southampton, Southampton SO17 1BJ, UK

‡ Environmental Change Unit, University of Oxford, Oxford OX1 3TB, UK

§ Hadley Centre for Climate Prediction and Research, The UK Meteorological Office, London Road, Bracknell RG12 2SY, UK

Assessments of the regional impacts of human-induced climate change on a wide range of social and environmental systems are fundamental for determining the appropriate policy responses to climate change^{1–3}. Yet regional-scale impact assessments are fraught with difficulties, such as the uncertainties of regional climate-change prediction⁴, the specification of appropriate environmental-response models⁵, and the interpretation of impact results in the context of future socio-economic and technological change⁶. The effects of such confounding factors on estimates of climate-change impacts have only been poorly explored^{3–7}. Here we use results from recent global climate simulations⁸ and two environmental response models^{9,10} to consider systematically the effects of natural climate variability (30-year timescales) and future climate-change uncertainties on river runoff and agricultural potential in Europe. We find that, for some regions, the impacts of human-induced climate change by 2050 will be undetectable relative to those due to natural multi-decadal climate variability. If misleading assessments of—and inappropriate adaptation strategies to—climate-change impacts are to be avoided, future studies should consider the impacts of natural multi-decadal climate variability alongside those of human-induced climate change.

Most climate-change impact studies follow a conventional methodology^{11,12} of simulating an environmental indicator—for example, runoff, crop yield, natural biome—first, under current climate, and then under a scenario of climate change. These

Table 1 Global and European climate changes for HadCM2 simulations.

	* $\Delta p\text{CO}_2$ (p.p.m.v.)	† $\Delta T_{\text{ANNglobal}}$ (°C)	† $\Delta T_{\text{ANNEurope}}$ (°C)	† ΔP_{DJF} Europe (%)	† ΔP_{JJA} Europe (%)
‡Control	334	±0.30	±0.24	±1.90	±3.64
\$GGA-1	515	1.97	2.57	8.48	−0.33
GGA-2	515	1.88	2.50	5.53	−2.25
GGA-3	515	1.93	2.05	9.37	−4.09
GGA-4	515	1.86	2.09	4.27	−3.33
GGA-mean	515	1.91	2.30	6.89	−2.51
\$GGd-1	435	1.42	1.70	6.85	−2.18
GGd-2	435	1.31	1.83	2.40	−0.11
GGd-3	435	1.34	1.48	7.79	−1.61
GGd-4	435	1.44	1.89	5.00	−0.27
GGd-mean	435	1.38	1.72	5.49	−1.04

* Actual CO_2 concentrations for GGA and GGd are estimated using the mix of greenhouse gases in, respectively, the IS92a and IS92d emissions scenarios³⁰, together with the total CO_2 -equivalent forcing used in the HadCM2 experiments.

† ANN, annual; DJF, boreal winter; JJA, boreal summer. Climate changes over Europe are calculated for land only.

‡ Control climate numbers are 2 σ model estimates of natural variability on 30-year timescales.

\$GGA and GGd are, respectively, greenhouse-gas-only forcing scenarios equivalent to a 1% and 0.5% yr^{-1} increase in CO_2 concentration from 1990 to 2100. Simulations 1 to 4 represent individual simulations in each ensemble, each with the same forcing but with different initial model conditions¹⁷. GGA and GGd numbers are model-simulated climate changes extracted for the period 2035–64. All changes are expressed with respect to the 1961–90 mean climate.

scenarios typically represent some future period, for example the year 2050, or some future atmospheric condition, for example twice the pre-industrial level of atmospheric CO₂. If the simulation model incorporates non-climate parameters (for example, fertilizer application) then these may also be perturbed to reflect a different future physical or technological environment. The difference between the two simulations is then an estimate of the impact of human-induced climate change, that is, the 'impact signal'.

Much of the work reported by the Intergovernmental Panel on Climate Change^{3,7} has followed this approach, and many estimates of climate-change damage functions used in economic analyses rely on these types of results^{13–16}. This methodology is extended here by exploring the effects on impact indicators of uncertainty in the baseline climate (that is, natural climate variability inducing 'impact noise') and uncertainty in the climate-change signal. The former is achieved by using multi-century simulations of unforced climate and the latter through the use of ensemble scenario simulations. A climate-change ensemble consists of a number of simulations made with the same climate model and identical external forcing, but with each simulation starting from slightly different initial conditions. Ensemble simulations therefore provide a representation of the 'unpredictability' of the climate system. This approach allows us to interpret simulated climate-change impacts in a probabilistic framework based on signal-to-noise ratios. Two impact indicators for Europe—runoff and water-limited wheat yield—are used to illustrate the method. We focus on changes in the mean, although our approach can also be applied to extreme events and we provide an example of this.

We use results from a recent set of climate-change experiments performed by the UK Hadley Centre with their coupled ocean–atmosphere model (HadCM2^{8,17}). These experiments include a long multi-century control simulation and four ensemble simulations, each ensemble being forced with a different anthropogenic forcing scenario. Seven independent 30-year climates were extracted from the first 240 years of the control simulation, along with 30-year climates for the periods 1961–90 and 2035–64 from two of the four ensembles: GGa and GGd. GGa represents a widely used forcing scenario of 1% yr⁻¹ growth in CO₂-equivalent concentrations for the next century, while GGd represents a low forcing scenario of

0.5% yr⁻¹ growth¹⁷ (see details in Table 1). These extracted model climates—both control and scenario—are converted into mean climate anomalies from the 1961–90 period of the respective model simulations, which are then added to the observed 1961–90 climate. Inter-annual and inter-daily climate variability is assumed to be unchanged in all cases, our concern in this study being to examine the relative impacts of multi-decadal (30-year timescale) climate variability. We treat the climates resulting from the control simulation as describing multi-decadal 'natural climate variability' (although model-simulated rather than observed) and those from the scenario simulations as describing 'climate change' by the year 2050. Other climate models would yield different results to those shown here, and a full range of forcing scenarios would be necessary for a comprehensive analysis of impact uncertainty; we illustrate one important aspect of uncertainty, however, using the selected data. The observed baseline climate data are monthly mean values for the 1961–90 period gridded at a 0.5° latitude/longitude resolution¹⁸. The model anomalies at a resolution of 2.5° latitude by 3.75° longitude were interpolated to the observed resolution using a gaussian space-filtering routine. More sophisticated 'down-scaling' methods exist¹⁹, but were not applied here.

Runoff was simulated at a spatial scale of 0.5° latitude/longitude, using a daily water balance model applied separately in each cell with soil and vegetation parameters derived from spatial databases⁹. The model does not simulate the direct effect of elevated CO₂ concentrations in reducing potential evapotranspiration, as this is assumed to be offset at the catchment scale by increased plant growth. The monthly input data are 'downscaled' to the daily scale, using a stochastic model that generates randomly distributed daily rainfall amounts constrained by the monthly total. The model simulates well the broad pattern of runoff across Europe, and has been used to assess the implications of climate change⁹.

There are three findings from the application of the hydrological model with the different climate data sets. First, the spatial patterns of runoff change resulting from climate change are different to those associated with natural climate variability. The climate change scenarios result in a strong north–south gradient of change⁹, while the runoff anomalies due to climate variability are less geographically structured. Second, the patterns of runoff change

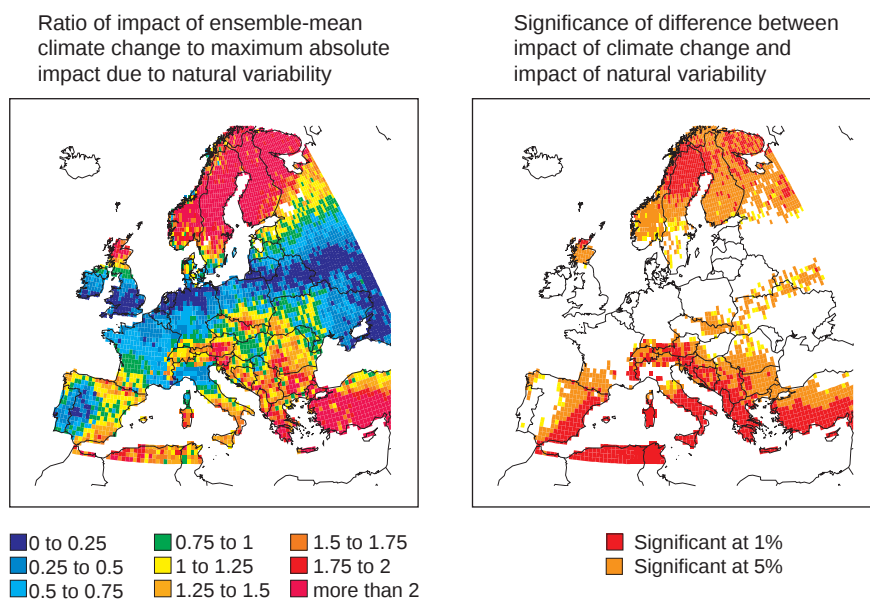


Figure 1 Relative impact of climate change on runoff. Change in mean annual runoff on a grid of 0.5° resolution across Europe due to climate change by 2050 under the GGa forcing scenario, expressed relative to the change in runoff due to natural climate variability. Left panel, absolute ensemble-mean per cent change in mean annual runoff by 2050 expressed as a ratio of the absolute maximum per

cent change in annual runoff caused by natural climate variability. Ensemble-mean runoff change is the average of the four runoff changes induced by the four ensemble climate changes. Right panel, significance of the difference in mean annual runoff induced by 2050 climate change ($n=4$) and by natural climate variability ($n=7$); two-tailed t-test.

are less consistent between the different natural climates (an average impact pattern correlation amongst the seven experiments of 0.21) than between the climate-change scenarios (average impact pattern correlations of 0.78 and 0.56 for GGa and GGd, respectively, with the lower correlations under GGd reflecting the weaker climate forcing of these simulations). Third, and most important at the catchment and water management scale, the impacts of multi-decadal climate variability may be greater than the impacts of climate change (Fig. 1). For this climate model and for these experiments, the impacts of climate change on mean runoff by 2050 are significantly greater than natural climate variability in northern and southern Europe, but are no different to those of natural climate variability across large parts of western and central Europe. Similar results were found using the value of the monthly runoff that was exceeded 90% of the time; this is a measure of low flow, calculated from the 30 individual years of simulated monthly runoff. This is because although the percentage impact of climate change on this extreme is greater than on the mean, the impact of multi-decadal natural climate variability on this measure is also greater. Different results may be obtained with other definitions of

extreme behaviour; where an increase in mean precipitation is associated with a greater relative increase in precipitation intensity, for example, the flood signal might expect to be strengthened.

Winter wheat yield was simulated at this same spatial resolution using the EuroWheat model¹⁰. EuroWheat is a simple mechanistic crop model that simulates the response of wheat development and yield to changes in climate (temperature, precipitation, solar radiation, relative humidity and wind speed) and atmospheric concentrations of carbon dioxide (through altered radiation use and water-use efficiencies). The model operates at two interacting time intervals: development and potential growth are determined on a daily time-step whilst water limitations to growth are computed on a monthly basis. EuroWheat has been shown to emulate the behaviour of widely used daily models (for example, AFRCWHEAT^{20,21}, CERES^{22,23}) under current and future climates at two sites in Europe¹⁰. It is also able to reproduce well the observed spatial pattern of wheat yield and has been used to investigate climate-change impacts on European wheat productivity^{10,24,25}.

Wheat productivity in Europe is highly sensitive to multi-decadal natural climate variability, with local (that is, 0.5° grid cell) changes in 30-year mean water-limited yield of up to $\pm 3 \text{ t ha}^{-1}$ (approximately $\pm 45\%$ of mean 1961–90 yield). When wheat yields are aggregated to the national scale, the absolute range of yield response to climate variability decreases, but still varies from 0.5 t ha^{-1} (8% of mean 1961–90 yield) in Italy to 1.3 t ha^{-1} (14.6%) in Romania (Fig. 2, top panel). Climate change alone—without considering effects on the plant of elevated CO_2 concentrations—only results in significant changes (in this case increases) in mean yield by 2050 for both GGa and GGd scenarios in three northern European countries: Finland, Germany and the Netherlands (Fig. 2, top panel). Elsewhere in Europe, climate-change impacts on mean wheat yield are indistinguishable from those due to natural climate variability, although Denmark, France and Italy experience significant yield increases under at least one of the two forcing scenarios. Including the effects of elevated CO_2 alters the signal-to-noise ratio for wheat yield substantially (Fig. 2, bottom panel). Ensemble-median national yields increase by between 1.2 (15%) and 2.9 t ha^{-1} (39%) for the GGa scenario and between 0.9 (9%) and 1.8 t ha^{-1} (29%) for the GGd scenario. All these increases are statistically significant. This high sensitivity of wheat yield to CO_2 concentration has been reported in many experimental and modelling studies^{26–28}, although whether such increases would be fully realised in the field and at large scales is less clear.

We believe that this is the first use in climate-change impact analysis of climate model results extracted both from long simulations of unforced climate and from ensemble scenario simulations. Conventional impact analyses implicitly assume that each scenario represents just the signal of climate change and that multi-decadal natural climate variability can be ignored. Our analysis shows that these assumptions are not correct. First, there is considerable variability in unforced 30-year climates and this natural climate variability can induce non-trivial responses in impact indicators. This conclusion is strengthened given the likelihood that the climate model simulations used here have underestimated natural climate variability²⁹. Second, climate-change scenarios derived from ensembles of climate model simulations, even when averaged over 30 years, contain differing magnitudes of both climate variability and climate change and this induces substantial uncertainty in the impact signal.

The particular patterns and magnitudes of change in runoff and wheat yield produced by our analysis may be specific to the climate and impact models used and to our choice of impact indicator (mean or extremes). Furthermore, the impact signal may be either over- or under-estimated, given that we have not simultaneously considered changes in variables such as land use, fertiliser application, pests, soil erosion or plant breeding. The way we interpret these changes, however, has wider validity and the results have a number of important general implications.

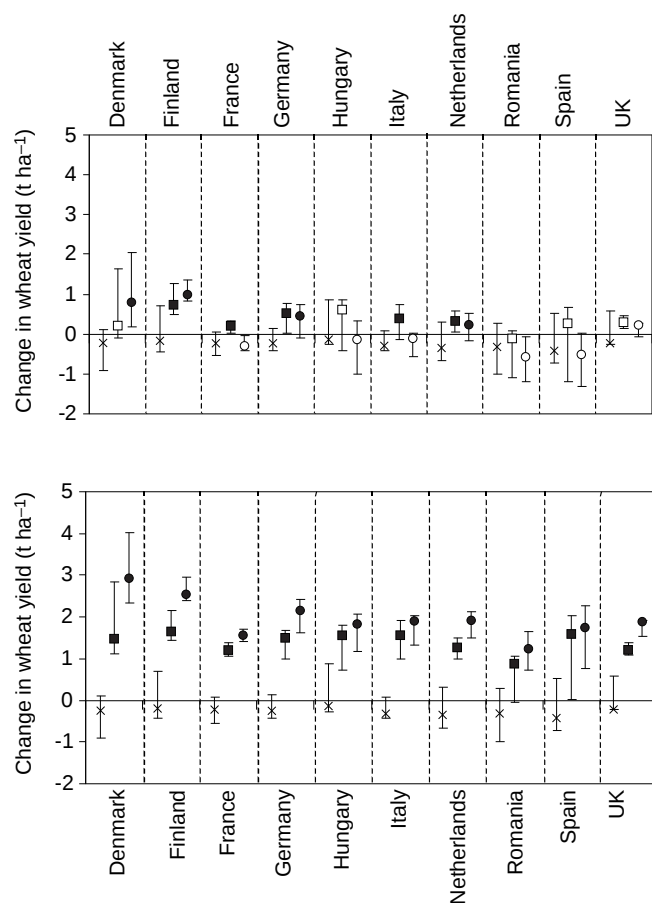


Figure 2 Relative changes in mean wheat yield at a national scale. Change in mean water-limited wheat yield (t ha^{-1} change from 1961–90 yield) for ten European countries due to natural climate variability (left bar; $n=7$) and due to climate and CO_2 change by 2050 under the GGd (middle bar; $n=4$) and GGa forcing scenarios (right bar; $n=4$). Top panel, atmospheric CO_2 concentration is held constant at 334 parts per million by volume (p.p.m.v.). Bottom panel, atmospheric CO_2 concentration increases to 435 p.p.m.v. for the GGd scenario and to 515 p.p.m.v. for the GGa scenario. Horizontal lines show maximum and minimum estimates in each case. Ensemble-median estimates are marked with an open square (GGd) or circle (GGa) where climate change causes a non-significant change in yield and with a filled square or circle where the change is significant at the 95% level from two-tailed t -test. Ensemble-median impacts are the median of the four impacts induced by the four ensemble-member climate changes.

First, in some sectors and for some regions human-induced climate change may not have as great an impact on natural resources as might multi-decadal natural climate variability. Comparing present resources only with those simulated under future climate change may exaggerate the importance of climate change by ignoring the impacts of natural variability on these time-scales: the estimated impacts may occur even in the absence of human-induced climate change. Second, the results suggest that in many areas it will be very difficult to detect the impact of climate change, even on a multi-decadal time scale; the different spatial patterns of climate change and climate-variability impact, suggest that detection is best undertaken by looking over a large geographic area. Third, adapting our management systems to withstand multi-decadal natural climate variability (adequately defined) may, in some sectors and for some regions, be a sufficient medium-term response to the prospect of climate change—although elsewhere it may not. Last, the results do not suggest that we can ignore the possibility that climate change will affect our natural resource base; what they do show is that some impacts of natural climate variability may be as great as, or greater than, the estimated impacts of human-induced climate change. This study shows that it is possible, and suggests that it is important, to compare the impacts of climate change alongside those of natural multi-decadal climate variability in order both to assess the importance of climate change and to help in the development of appropriate adaptation strategies. □

Received 30 April; accepted 30 November 1998.

1. Parry, M. L., Carter, T. R. & Hulme, M. What is a dangerous climate change? *Glob. Environ. Change*, **6**, 1–6 (1996).
2. Wigley, T. M. L., Richels, R. & Edmonds, J. A. Economic and environmental choices in the stabilization of atmospheric CO₂ concentrations. *Nature* **379**, 240–243 (1996).
3. IPCC (eds Watson, R. T., Zinyowera, M. C., Moss, R. H. & Dokken, D. J.) *The Regional Impacts of Climate Change: an Assessment of Vulnerability* (Cambridge Univ. Press, 1998).
4. Kittel, T. G. F., Giorgi, F. & Mehl, G. A. Intercomparison of regional biases and doubled CO₂—sensitivity of coupled atmosphere—ocean general circulation model experiments. *Clim. Dyn.* **14**, 1–15 (1998).
5. Shugart, H. H. & Smith, T. M. A review of forest patch models and their applications to global change research. *Clim. Change* **34**, 131–153 (1996).
6. Dowlatabadi, H. Assessing the health impacts of climate change. *Clim. Change* **35**, 137–144 (1997).
7. Watson, R. T., Zinyowera, M. C. & Moss, R. H. (eds) *Climate Change 1995: Impacts, Adaptations and Mitigation of Climate Change: Scientific-Technical Analyses* (Cambridge Univ. Press, 1996).
8. Mitchell, J. F. B. & Johns, T. C. On the modification of global warming by sulphate aerosols. *J. Clim.* **10**, 245–267 (1997).
9. Arnell, N. W. The effect of climate change on hydrological regimes in Europe: a continental perspective. *Glob. Environ. Change* **9**, 5–23 (1999).
10. Harrison, P. A. & Butterfield, R. E. Effects of climate change on Europe-wide winter wheat and sunflower productivity. *Clim. Res.* **7**, 225–241 (1996).
11. Carter, T. R., Parry, M. L., Harasawa, H. & Nishioka, S. *IPCC Technical Guidelines for Assessing Climate Change Impacts and Adaptations* (UCL/CGER, London/Tsukuba, 1994).
12. Parry, M. L. & Carter, T. *Climate Impact and Adaptation Assessment* (Earthscan, London, 1998).
13. Hope, C., Anderson, J. & Wenman, P. Policy analysis of the greenhouse effect. *Energy Policy* **21**, 327–338 (1993).
14. Nordhaus, W. D. Optimal greenhouse gas reductions & tax policy in the 'DICE' model. *Am. Econ. Rev. Pap. Proc.* **83**, 313–317 (1993).
15. Fankhauser, S. *Valuing Climate Change* (Earthscan, London, 1995).
16. Tol, R. S. J. The damage costs of climate change: toward more comprehensive calculations. *Environ. Resource Econ.* **5**, 353–374 (1995).
17. Mitchell, J. F. B., Johns, T. C., Eagles, M., Ingram, W. J. & Davis, R. A. Towards the construction of climate change scenarios. *Clim. Change* (in the press).
18. Hulme, M. *et al.* Construction of a 1961–90 climatology for Europe for climate change impacts and modelling applications. *Int. J. Climatol.* **15**, 1333–1363 (1995).
19. Wilby, R. & Wigley, T. M. L. Downscaling general circulation model output: a review of methods and limitations. *Prog. Phys. Geogr.* **21**, 530–548 (1997).
20. Ritchie, J. & Otter, S. in *ARS Wheat Yield Project* (ed. Willis, W. O.) 159–175 (Dept of Agriculture, Agriculture Research Service, ARS-38, Washington DC, 1985).
21. Godwin, D., Ritchie, J., Singh, U. & Hunt, L. *A User's Guide to CERES-Wheat - V2.10* (Simulation manual IFDC-SM-2, Int. Fertilizer Development Center, Muscle Shoals, AL, 1990).
22. Weir, A. H., Bragg, P. L., Porter, J. R. & Rayner, J. H. A winter wheat crop simulation model without water or nutrient limitations. *J. Agric. Sci.* **102**, 371–382 (1984).
23. Porter, J. R. AFRCWHEAT2: a model of the growth and development of wheat incorporating responses to water and nitrogen. *Eur. J. Agronomy* **2**, 69–82 (1993).
24. Harrison, P. A. Modelling the effects of climate change on wheat productivity in Europe. *Aspects Appl. Biol.* **45**, 41–48 (1996).
25. Harrison, P. A., Butterfield, R. E. & Gawith, M. J. in *Climate Change and Agriculture in Europe: Assessment of Impacts and Adaptations* (eds Harrison, P. A., Butterfield, R. E. & Downing, T. E.) 370–379 (Res. Rep. No. 9, Environmental Change Unit, Univ. Oxford, 1995).
26. Cure, J. D. & Acock, B. Crop responses to carbon dioxide doubling: a literature survey. *Agric. Forest Meteorol.* **38**, 127–145 (1986).
27. Kimball, B. A. *et al.* Productivity and water-use of winter wheat under free-air CO₂ enrichment. *Glob. Change Biol.* **1**, 429–442 (1995).
28. Batts, G. R. *et al.* Yield and partitioning in crops of contrasting cultivars of winter wheat in response to CO₂ and temperature in field studies using temperature gradient tunnels. *J. Agric. Sci.* **130**, 17–27 (1998).

29. Tett, S. F. B., Johns, T. C. & Mitchell, J. F. B. Global and regional variability in a coupled AOGCM. *Clim. Dyn.* **13**, 303–323 (1997).
30. Leggett, J., Pepper, W. J. & Swart, R. J. in *Climate Change 1992: the Supplementary Report to the IPCC Scientific Assessment* (eds Houghton, J. T., Callander, B. A. & Varney, S. K.) 75–95 (Cambridge Univ. Press, 1992).

Acknowledgements. Model data were obtained through the Climate Impacts LINK Project. This work was supported by DGXII of the Commission of the European Community.

Correspondence and requests for materials should be addressed to M.H. (e-mail: m.hulme@uea.ac.uk).

The origin of spinifex texture in komatiites

Mark Shore & Anthony D. Fowler

Ottawa-Carleton Geoscience Centre and Department of Earth Sciences,
University of Ottawa, 140 Louis Pasteur, PO Box 450, Station A, Ottawa,
Ontario K1N 6N5, Canada

Komatiites are high-temperature, fluid, magnesium-rich lavas typically of Archaean age. A striking characteristic feature of such lavas is 'spinifex' texture—plate-like crystals of olivine ((Mg,Fe)₂SiO₄), millimetres to decimetres long, in a fine-grained matrix of spherulitic clinopyroxene (Ca(Mg,Fe,Al)(Si,Al)₂O₆), dendritic chromite ((Mg,Fe)(Cr,Al,Fe)₂O₄) and altered glass^{1–4}. Sheaves of olivine crystals can reach lengths exceeding one metre, even in komatiite flows less than 10 metres thick, in sharp contrast to the millimetre-scale post-eruption growth of crystals in more common volcanic rocks. Crystal growth of this magnitude might be a consequence of the high content of the constituent elements of olivine in komatiitic liquid, combined with the low viscosity and high chemical diffusivity of the lavas. But flows lacking spinifex texture are not uncommon, and those with such texture often contain substantial amounts of submillimetre olivine crystals of unremarkable appearance, so chemical considerations alone do not appear to provide a sufficient explanation. Here we present evidence that spinifex texture develops as a result of large thermal gradients, coupled with conductive and radiative heat transfer within olivine crystals fixed in the cool upper layers of the lava flows. This mode of growth has features in common with the high-temperature techniques used to grow large synthetic single crystals, but is rarely considered in geological contexts.

Many komatiites, including the exceptionally fresh and well exposed flows at Pyke hill in Munro township, northeastern Ontario, erupted as subaqueous lava flows^{1,4}. The thermal effect of seawater infiltration into the fractured upper crust of the flows has not previously been considered in komatiite cooling models^{5,6}. Thermal contraction of the solidified sheet-like lava flows (below ~1,000 °C) caused extensive fracturing on a centimetre to decimetre scale (Fig. 1). The permeability of a rock body can be estimated from the number and mean width of its fractures⁷; that of the illustrated Pyke hill lava tube is ~10⁻¹¹ m⁻². This is close to the calculated bulk permeability of young oceanic crust (~10⁻¹² m⁻²)⁸; by contrast, unfractured igneous rocks have typical *in situ* permeabilities of the order of 10⁻¹⁶ to 10⁻¹⁸ m⁻². Thus we expect a substantial heat flux to have been transported by water circulation through cracks. Semiquantitative measurements of hydrothermal heat fluxes from basalt flows include ~40 kW m⁻² at Heimaey⁹, 100 kW m⁻² (with transient fluxes approaching 1 MW m⁻²) at Kilauea¹⁰, and ~1 MW m⁻² at Vatnajökull¹¹ (estimated from data in that reference). Such heat transfer rates cannot be achieved by conduction through more than a few centimetres of solid rock. Because fracturing leads to further ingress of sea water, hydrothermal cooling/cracking fronts are probably self-propagating (at least in thin flows) and would migrate rapidly downwards. Observations^{9–11} and modelling¹² of subaqueous basalt flows have demonstrated that cooling fronts could move at rates as high as several decimetres per hour, resulting in thermal gradients >10⁴ K m⁻¹ and

much faster cooling of flow interiors than would occur via purely conductive cooling.

Rapid cooling, coupled with the absence of pre-existing nuclei of clinopyroxene and plagioclase (komatiite lavas were commonly superheated with respect to these minerals³), can explain the spherulitic habit and aluminium-rich composition of clinopyroxene (high pressure, high temperature and rapid growth are some factors promoting coupled substitution of 2 Al for Mg + Si in clinopyroxene^{13,14}), as well as the delayed crystallization of plagioclase in many komatiites. However, rapid cooling and high thermal gradients offer only a partial explanation for spinifex texture, as komatiite flows typically contain a substantial proportion of millimetre-sized olivine crystals of similar morphology to those of basalt and other common rock types.

Textural evidence, including the preferred orientation of the olivine sheaves perpendicular to cooling contacts and their thickening away from the flow surface (Fig. 2), indicates that downward growth of olivine was a competitive process^{2,3}. At the microscopic scale, we observe (Fig. 3) that the growth of plate-like olivine crystals was commonly blocked by earlier-formed crystals. Growth of the later crystals proceeded only on faces away from the blocking crystal owing to a lack of available space and material (primarily Mg).

The nucleation and growth of olivine occur readily in ultramafic liquids^{15,16}, to the point that it is difficult to quench even strongly superheated liquids to crystal-free glasses. Chilled margins of Pyke hill flows, originally glassy, now contain sparse millimetre-sized euhedral olivine crystals as well as hundreds to thousands of micrometre-sized olivine crystallites per cubic centimetre (ref. 4). Thus, even minor differences in olivine growth rates, due to favourable orientation with respect to thermal gradients, gravity or crystallographic axes, could produce a strongly anisotropic fabric over the hours to days required to solidify a thin komatiite flow.

An important factor is the large difference in the thermal properties of olivine and molten komatiite at the temperatures of interest (~1,300–1,550 °C). The mean lattice (or phonon) thermal conductivity k_L of olivine is estimated to be 2–3 W m⁻¹ K⁻¹, as extrapolated from lower-temperature measurements ($T = 20$ –1,200 °C)^{17–20} assuming $k_L = 1/(a + bT)$, where a and b are constants. This relationship is not exact, but is reasonably accurate at ambient to high temperatures for olivine and other ionic crystals with moderately complex structures¹⁹. Inconsistencies between different data sets are largely due to the variable effects of crystallographic orientation in single-crystal studies, grain boundaries in polycrystalline material, geometrically dependent radiative heat transfer, oxidation/reduction of samples at high temperature, and overall experimental difficulties of measurements at high pressures.

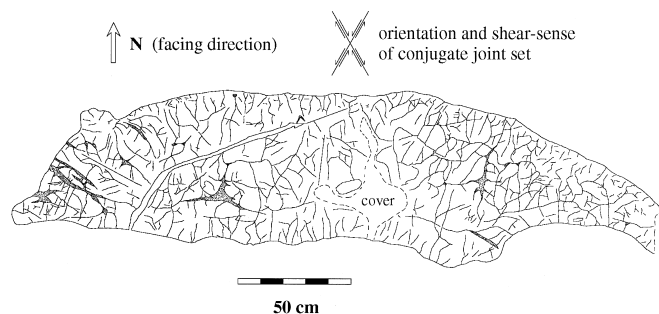


Figure 1 Fractures within a komatiite lava tube, Pyke hill, Ontario. Although the tube lacks spinifex texture, nearby spinifex-textured flows have similar fracture densities. Most fractures are unrelated to the conjugate set (orientation shown above) associated with east–west regional folding. Measurements of average crack width (~0.5 mm) and fracture length per unit area (28 m⁻¹) lead to an estimated permeability of 10⁻¹¹ m⁻².

Because of the high temperature of molten komatiites, we must also consider radiative heat transfer in olivine, which has previously been studied as a mechanism of heat flow within the Earth's mantle. The radiative thermal conductivity (k_R) within an optically thick body (such that heat transfer can be considered a diffusive process) at thermal equilibrium, with temperature varying in one direction, is described by the expression²¹

$$k_R = \frac{20\sigma T^3}{\pi^4} \int_0^\infty l_x \frac{e^x x^4 n_x^2}{(e^x - 1)^2} dx \quad (1)$$

where $x = h\nu/(kT)$, l_x and n_x are the photon mean free path and refractive index at frequency ν and mean temperature T , h is Planck's constant (6.626×10^{-34} J s), k is Boltzmann's constant (1.38×10^{-23} J K⁻¹), and σ is the Stefan–Boltzmann constant (5.67×10^{-8} W m⁻² K⁻⁴). Where scattering is minimal, as in a single crystal, the mean free path approximately equals the reciprocal of the absorption coefficient. Calculations²¹ using equation (1) and the measured visible to mid-infrared absorption spectra of magnesian olivine (Fe/(Mg + Fe) = 0.08, similar to komatiitic olivines) show that the effective radiative thermal conductivity within single crystals is 2.0–2.2 W m⁻¹ K⁻¹ at 1,400 °C. Although blackbody emissivity increases by 40% from 1,400 to 1,550 °C, temperature-induced broadening of absorption bands largely offsets the increased emissivity²¹. Assuming an effective radiative thermal conductivity for komatiitic olivine of ~2 W m⁻¹ K⁻¹, we arrive at a total thermal conductivity of 4–5 W m⁻¹ K⁻¹ for olivine, calculated as the sum of the lattice and effective radiative thermal conductivities. Similar results have been obtained by direct measurements of combined lattice/radiative conductivity of single-crystal specimens several millimetres in size^{17,19}.

Owing to experimental difficulties, measurements of the thermal conductivity of molten silicates vary greatly. We assume a value of ~1 W m⁻¹ K⁻¹ for the thermal conductivity (including transmission of thermal radiation) of molten komatiite, consistent with experimental measurements on iron-bearing silicate liquids (0.5–1.5 W m⁻¹ K⁻¹)^{22,23}, as other measurements on iron-free silicate melts

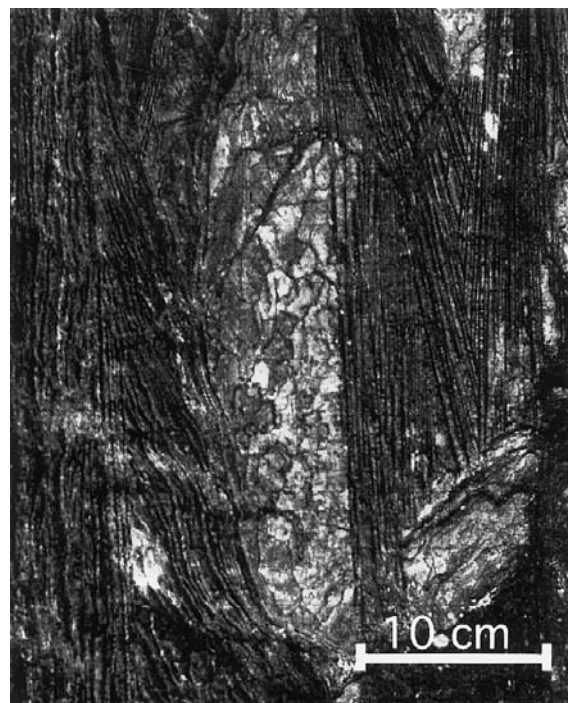


Figure 2 Well-developed spinifex texture in a komatiite flow at Pyke hill. Large sheaves of subparallel olivine crystals broaden towards the base of the flow. The plate-like olivine crystals in the central sheaf have their largest faces parallel to the outcrop surface.

showing very low thermal conductivities at high temperatures ($0.04 \text{ W m}^{-1} \text{ K}^{-1}$ at $1,600^\circ\text{C}$) may be flawed²⁴. Even using the higher value for the thermal conductivity of the molten silicate, a robust conclusion is that the thermal conductivity of olivine is approximately three to five times greater than that of its growth liquid.

We therefore propose that olivine, rooted in the hydrothermally cooled upper layer of flows, grew by (thermally) constrained crystallization²⁵; this is a process wherein crystals grow into (and transport heat from) a hotter liquid. This differs from most magmatic crystallization wherein the latent heat of crystallization is dissipated in cooler liquid surrounding the crystals. In effect, despite the release of latent heat, the growing olivine crystals would act as heat sinks for the surrounding komatiitic liquid, not as heat sources. In contrast, olivine crystals freely suspended in liquid (as in a flowing komatiite lava) would grow normally, and would not be expected to show unusual morphological features.

The most pronounced effect of the different thermal and optical properties of liquid and olivine would occur at the leading tip of a growing crystal. The result is to increase local thermal gradients and make the crystal–melt interface more convex (that is, sharpen crystal tips), as shown by observations²⁶ and finite-element modelling^{27,28} of Czocharski and Bridgman growth of oxide single crystals.

We now examine evidence supporting this mode of growth, considering first the crystallographic properties of olivine in greater detail. The thermal conductivity of olivine is strongly anisotropic, with values (at ambient temperature) of $5.9 (\pm 0.1)$, $3.4 (\pm 0.06)$ and $5.1 (\pm 0.1) \text{ W m}^{-1} \text{ K}^{-1}$ parallel to the *a*, *b* and *c* crystallographic axes

(space group *Pbnm*), respectively²⁹. This anisotropy persists at higher temperatures (M. D. Collins, personal communication). In addition, olivine is strongly pleochroic (its optical absorption varies with crystallographic orientation of light) in the near-infrared range from 0.8 to $1.2 \mu\text{m}$ (refs 21, 30), the spectral region of maximum blackbody emission from a partially molten komatiite. At these wavelengths, absorption is lowest for non-polarized light propagating parallel to the *a* crystallographic axis. The total thermal conductivity at high temperatures is expected to vary as $a > c \gg b$, with conductivity parallel to *a* being 10–20% greater than that parallel to *c*. Downward growth parallel to the *a* axis is predicted if heat transfer from the leading tips of crystals is a dominant growth-rate factor.

Accordingly, we examined olivine within coarse spinifex-textured lavas from Pyke hill for evidence of a preferred crystallographic orientation (in addition to the previously noted flattening perpendicular to the *b* axis²). Orientated thin sections were cut from 14 different olivine sheaves that were perpendicular ($\pm 20^\circ$) to the upper surface of flows. In every case, the olivine crystals were orientated with *a* perpendicular to the flow tops, parallel to the inferred maximum thermal gradient. The same orientation has been found in synthesized comb-layered olivine, a texture also attributed to constrained crystallization³¹.

This growth mechanism might be applicable to the vertically orientated needle-like clinopyroxene crystals found in some komatiitic basalts. But the lower temperatures involved ($\sim 1,150$ – $1,300^\circ\text{C}$) would reduce the role of radiative heat transfer, while the greater effect of superheating on clinopyroxene nucleation (versus that of olivine) and the role of the polymerized chain structure of pyroxene parallel to the *c* axis are complicating factors. In addition, the requisite high-temperature thermal conductivity and infrared spectral data are not available.

Thermally constrained crystallization of olivine within spinifex-textured layers of komatiite flows would have been favoured by specific conditions present (kinetically easy growth of olivine, elevated temperatures, large and consistently orientated thermal gradients, and significant differences in the thermal transport properties of olivine and molten komatiite). Preserved textural and mineralogical evidence is consistent with this geologically unusual mode of growth. Other processes that probably occurred within cooling flows (for example, compositionally and thermally driven melt convection^{5,6}) may have been of less importance to the development of spinifex texture. □



Figure 3 Cross-section of plate-like olivine crystals (largely replaced by serpentine and magnetite) in a spinifex-textured komatiite. The crystal to the left has interfered with the growth of its neighbour, which has formed a series of stepped-out segments to the right. Gently curved olivine crystals are common in komatiites. In such crystals the crystallographic orientation does not vary; rather, the curvature arises from numerous small growth offsets, all in the same direction.

Received 16 February; accepted 14 December 1998.

- Pyke, D. R., Naldrett, A. J. & Eckstrand, O. R. Archean ultramafic flows in Munro Township, Ontario. *Bull. Geol. Soc. Am.* **84**, 955–978 (1973).
- Donaldson, C. H. in *Komatiites* (eds Arndt, N. T. & Nisbet, E. G.) 213–244 (Allen & Unwin, London, 1982).
- Arndt, N. T. in *Archean Crustal Evolution* (ed. Condie, K. C.) 11–44 (Elsevier, Amsterdam, 1994).
- Shore, M. *Cooling and Crystallization of Komatiite Flows* Thesis, Univ. Ottawa (1996).
- Huppert, H. E. & Sparks, R. S. J. Komatiites I: Eruption and flow. *J. Petrol.* **26**, 694–725 (1985).
- Turner, J. S., Huppert, H. E. & Sparks, R. S. J. Komatiites II: Experimental and theoretical investigations of post-emplacement cooling and crystallization. *J. Petrol.* **27**, 397–437 (1986).
- Nehlig, P. Fracture and permeability analysis in magma-hydrothermal transition zones in the Samail ophiolite (Oman). *J. Geophys. Res.* **99**, 589–601 (1994).
- Davis, E. E., Chapman, D. S. & Forster, C. B. Observations concerning the vigor of hydrothermal circulation in young oceanic crust. *J. Geophys. Res.* **101**, 2927–2942 (1996).
- Björnsson, H., Björnsson, S. & Sigurgeirsson, Th. Penetration of water into hot rock boundaries of magma at Grimsvötn. *Nature* **295**, 580–581 (1982).
- Hardee, H. C., Dunn, J. C. & Hills, R. G. Probing the melt zone of Kilauea Iki lava lake, Kilauea volcano, Hawaii. *Geophys. Res. Lett.* **8**, 1211–1214 (1981).
- Gudmundsson, M. T., Sigmundsson, F. & Björnsson, H. Ice–volcano interaction of the 1996 Gjalp subglacial eruption, Vatnajökull, Iceland. *Nature* **389**, 954–957 (1997).
- Lowell, R. P. & Germanovich, L. N. On the temporal evolution of high-temperature hydrothermal systems at ocean ridge crests. *J. Geophys. Res.* **99**, 565–575 (1994).
- Kinzel, R. T. & Grove, T. L. Crystallization and differentiation of Archean komatiite lavas from northeast Ontario: phase equilibrium and kinetic studies. *Am. Mineral.* **70**, 40–51 (1985).
- Parman, S. W., Dann, J. C., Grove, T. L. & de Wit, M. J. Emplacement conditions of komatiite magmas from the 3.49 Ga Komati Formation, Barberton Greenstone Belt, South Africa. *Earth Planet. Science Lett.* **150**, 303–323 (1997).
- Donaldson, C. H. An experimental investigation of olivine morphology. *Contrib. Mineral. Petrol.* **57**, 187–213 (1976).
- Donaldson, C. H. An experimental investigation of the delay in nucleation of olivine in mafic magmas. *Contrib. Mineral. Petrol.* **69**, 21–32 (1979).
- Kanamori, H., Fujii, N. & Mizutani, H. Thermal diffusivity measurement of rock-forming minerals from 300° to 1100°K . *J. Geophys. Res.* **73**, 595–605 (1968).

18. Fujisawa, H. *et al.* Thermal diffusivity of Mg_2SiO_4 , Fe_2SiO_4 , and NaCl at high pressures and temperatures. *J. Geophys. Res.* **73**, 4727–4733 (1968).
19. Schatz, J. F. & Simmons, G. Thermal conductivity of earth materials at high temperatures. *J. Geophys. Res.* **77**, 6966–6983 (1972).
20. Katsura, T. Thermal diffusivity of olivine under upper mantle conditions. *Geophys. J. Int.* **122**, 63–69 (1995).
21. Shankland, T. J., Nitsan, U. & Duba, A. G. Optical absorption and radiative heat transport in olivine at high temperature. *J. Geophys. Res.* **84**, 1603–1610 (1979).
22. Murase, T. & McBirney, A. R. Properties of some common igneous rocks and their melts at high temperatures. *Bull. Geol. Soc. Am.* **84**, 3563–3592 (1973).
23. Büttner, R., Zimanowski, B., Blumm, J. & Hagemann, L. Thermal conductivity of a volcanic rock material (olivine-melilitite) in the temperature range between 288 and 1470 K. *J. Volcanol. Geotherm. Res.* **80**, 293–302 (1998).
24. Shore, M. Comment on “Experimental determination of the thermal conductivity of molten $CaMgSi_2O_6$ and the transport of heat through magmas”. *J. Geophys. Res.* **100**, 22401–22402 (1995).
25. Tiller, W. A. *The Science of Crystallization: Macroscopic Phenomena and Defect Generation* (Cambridge Univ. Press, 1991).
26. Cockayne, B., Chesswas, M. & Gasson, D. B. Facetting and optical perfection in Czochralski grown garnets and rubies. *J. Mater. Sci.* **4**, 450–456 (1969).
27. Brandon, S. & Derby, J. J. Heat transfer in vertical Bridgman growth of oxides: effects of conduction, convection, and internal radiation. *J. Cryst. Growth* **121**, 473–494 (1992).
28. Tsukada, T., Kakinoki, K. & Hozawa, M. Effect of internal radiation within crystal and melt on Czochralski crystal growth of oxide. *Int. J. Heat Mass Transfer* **38**, 2707–2714 (1995).
29. Choi, M., Brown, J. M. & Slutsky, L. J. Thermal diffusivity of mantle minerals. *Phys. Chem. Miner.* **23**, 470–475 (1996).
30. Burns, R. G. *Mineralogical Applications of Crystal Field Theory* 2nd edn (Cambridge Univ. Press, 1993).
31. Donaldson, C. H. Laboratory duplication of comb layering in the Rhum pluton. *Mineral. Mag.* **41**, 323–336 (1977).

Acknowledgements. We thank A. M. Hofmeister for comments on the manuscript. This work was supported by NSERC.

Correspondence and requests for materials should be addressed to A.D.F. (e-mail: aforwer@uottawa.ca).

A ferric-chelate reductase for iron uptake from soils

Nigel J. Robinson*, Catherine M. Procter*, Erin L. Connolly† & Mary Lou Guerinot†

* Department of Biochemistry and Genetics, The Medical School, University of Newcastle, Newcastle NE2 4HH, UK

† Department of Biological Sciences, Dartmouth College, Hanover, New Hampshire 03755-3576, USA

Iron deficiency afflicts more than three billion people worldwide¹, and plants are the principal source of iron in most diets. Low availability of iron often limits plant growth because iron forms insoluble ferric oxides, leaving only a small, organically complexed fraction in soil solutions². The enzyme ferric-chelate reductase is required for most plants to acquire soluble iron. Here we report the isolation of the *FRO2* gene, which is expressed in iron-deficient roots of *Arabidopsis*. *FRO2* belongs to a superfamily of flavocytochromes that transport electrons across membranes. It possesses intramembranous binding sites for haem and cytoplasmic binding sites for nucleotide cofactors that donate and transfer electrons. We show that *FRO2* is allelic to the *frd1* mutations that impair the activity of ferric-chelate reductase³. There is a nonsense mutation within the first exon of *FRO2* in *frd1-1* and a missense mutation within *FRO2* in *frd1-3*. Introduction of functional *FRO2* complements the *frd1-1* phenotype in transgenic plants. The isolation of *FRO2* has implications for the generation of crops with improved nutritional quality and increased growth in iron-deficient soils.

Arabidopsis is an ‘iron-efficient’ plant, able to acquire iron from soils of low iron availability, but the genetic basis of iron efficiency is unclear. In such plants, the activity of ferric-chelate reductase at the plasma membrane of root epidermal cells increases when iron is deficient^{4,5}. Organic compounds of either plant or microbial origin that retain Fe^{3+} in the soil solution have a lesser affinity for Fe^{2+} . It is therefore likely that ferric-chelate reductase activity releases iron from organic compounds, generating free iron for uptake⁶. Because the ferric-chelate reductase of plant roots has some functional similarity to the human phagocytic NADPH oxidase gp91phox

(ref. 7) and yeast ferric-chelate reductases such as FRE1 (ref. 8) and FRP1 (ref. 9), we speculated that it may share elements of structural similarity with these enzymes. The enzymes are involved in the transfer of electrons from cytosolic donors to FAD and then, through two consecutive haem groups, to single electron acceptors on the opposing face of a membrane. The haem groups are coordinated to four conserved histidine residues located on two transmembrane α -helices¹⁰. However, gp91phox and the yeast ferric-chelate reductases have in common only a few short sequence motifs, most notably associated with the cofactor-binding sites.

Two approaches were initiated to obtain a ferric-chelate reductase gene: generating reductase-deficient plant mutants³, and cloning candidate plant DNA sequences. Degenerate oligonucleotide primers were designed to anneal to sequences encoding the only tetrapeptide motif (the FAD-binding site) common to gp91phox and the yeast ferric-chelate reductases, and to a second partly conserved motif associated with the NADPH-binding sites. Several polymerase chain reaction (PCR) products were amplified from *Arabidopsis* genomic DNA using a low annealing temperature. Seven products were cloned and sequenced and one, which encodes a serine-rich polypeptide, was selected for further analysis because FRP1 is serine rich in the region between the cofactor-binding sites. This PCR product was used to screen genomic and complementary DNA libraries, and a genomic fragment containing two closely related genes, designated *FRO1* and *FRO2*, was fully sequenced (Fig. 1a). The *FRO2* intron–exon boundaries were identified from the cDNA sequence; the *FRO1* intron–exon boundaries were deduced by analogy to *FRO2* by using predictions¹¹ of 5′ donor and 3′ acceptor sites and by reverse transcriptase PCR (RT-PCR). *FRO2* and *FRO1* have 61.9% sequence identity (90.5% similarity). Flanking the 3′ end of *FRO1* is a synvergently transcribed gene encoding a cytochrome P450, designated *CYP86A4* (Fig. 1a), followed by a convergently transcribed unknown gene with similarity to the expressed sequence tags Z17619, N96903 and H76348. Flanking the 5′ end of *FRO2* is a synvergently transcribed gene related to *MKP4* that encodes a deduced mitogen-activated protein (MAP) kinase (not shown on Fig. 1a).

The sequence and predicted structure of *FRO2* is shown in Fig. 1. *FRO2* is composed of 725 amino acids, has a predicted pI of 9.37 and relative molecular mass (M_r) of 81.50, although the presence of glycosylation motifs suggest that the native protein may have a greater M_r . *FRO2* contains sequences identical to the FAD-binding site and the conserved region adjacent to the NADPH-binding site of FRE1 (underlined in Fig. 1b) but no other identical sequence motifs of an equivalent length. However, predictions of secondary structure indicate further similarity. The amino-terminal regions of the yeast ferric-chelate reductases and gp91phox form several transmembrane α -helices that are apparent in hydrophobicity plots. Six hydrophobic domains are identified within the N-terminal regions of *FRO2* along with two carboxy-terminal hydrophobic domains, all of which are predicted¹² to form transmembrane α -helices. This suggests that an intracellular region of *FRO2*, containing the deduced cofactor-binding sites, is anchored at both ends by membrane-spanning regions (Fig. 1c). In contrast, the gp91phox and yeast ferric-chelate reductase cofactor-binding sites and associated C-terminal regions are all predicted to be cytosolic. To form an active complex, gp91phox requires a second membrane protein, p22phox. The total number of predicted transmembrane α -helices in *FRO2* is equivalent to the number in gp91phox (six) and p22phox (two) combined, and it is feasible that the C-terminal domain, including the two additional transmembrane α -helices in *FRO2*, performs an analogous function to p22phox (p22phox shows 10% identity and 50% sequence similarity to the C-terminal domain of *FRO2*). Two pairs of histidine residues in *FRO2* that lie on two predicted, similarly orientated, transmembrane α -helices are in equivalent locations (Fig. 1b) to the histidine residues in FRE1 and gp91phox that coordinate haem¹⁰ (Fig. 1c).

Arabidopsis frd1 mutants do not induce ferric-chelate reductase activity, so they do not translocate ^{55}Fe to shoots when roots are presented with $^{55}\text{Fe}^{3+}$ -chelates³. The *frd1* mutations map towards the top of chromosome 1, 0 centiMorgans from markers PVV4 (ATHACS) and 19 centiMorgans from NCCI (ref. 3). Partial DNA sequences from the ends of the bacterial artificial chromosome clones F9C19 and T1617 (GenBank accession numbers B11604 and B09869, respectively) show identity to sections of the MAP kinase gene and *CYP86A4* that flank *FRO2* and *FRO1*. F9C19 and T1617 cross-hybridize with the yeast artificial chromosome clones yUP20D1, yUP12D7 and CIC3H3, which contain sequences that also map near to PVV4 (ref. 13). These data indicate that the region shown in Fig. 1a is located on chromosome 1 at, or close to, the *frd1* mutations. Sequencing both cDNA and genomic DNA from allelic mutants *frd1-1* and *frd1-3* revealed non-synonymous nucleotide substitutions within *FRO2*: a nonsense mutation in the first exon of *frd1-1* and the substitution of a threonine codon with a methionine codon in *frd1-3* (Fig. 2a). The amino-acid substitution in *frd1-3* lies within the deduced tetrapeptide FAD-binding site. Relative to the

wild type, *frd1-1* mutants had impaired growth on media with no added iron (Fig. 2b) but equivalent growth on media supplemented with 100 μM Fe^{3+} EDTA (Fig. 2c).

We generated transgenic *frd1-1* plants that carry a genomic DNA fragment including *FRO2* and 0.6 kilobase (kb) of sequence 5' of the *FRO2* ATG start codon. Low-iron-inducible root-surface ferric-chelate reductase activity was restored in these plants (Fig. 3), confirming that the *frd1* phenotypes are brought about by the observed mutations in *FRO2*. An equivalent restoration of ferric-chelate reductase activity was observed in three independent transformed lines.

Semiquantitative RT-PCR reveals that *FRO2* transcripts are more abundant than *FRO1* transcripts in roots (data not shown) and that *FRO2* transcripts accumulate in response to iron deficiency (Fig. 4a). This suggests that increased ferric-chelate reductase activity under these conditions involves the synthesis of the enzyme, rather than merely the activation of pre-existing reductase. The amount of *FRO2* transcript is reduced in *frd1-1* and increased in *frd1-3* (Fig. 4b). Premature termination of translation of *FRO2* in *frd1-1*

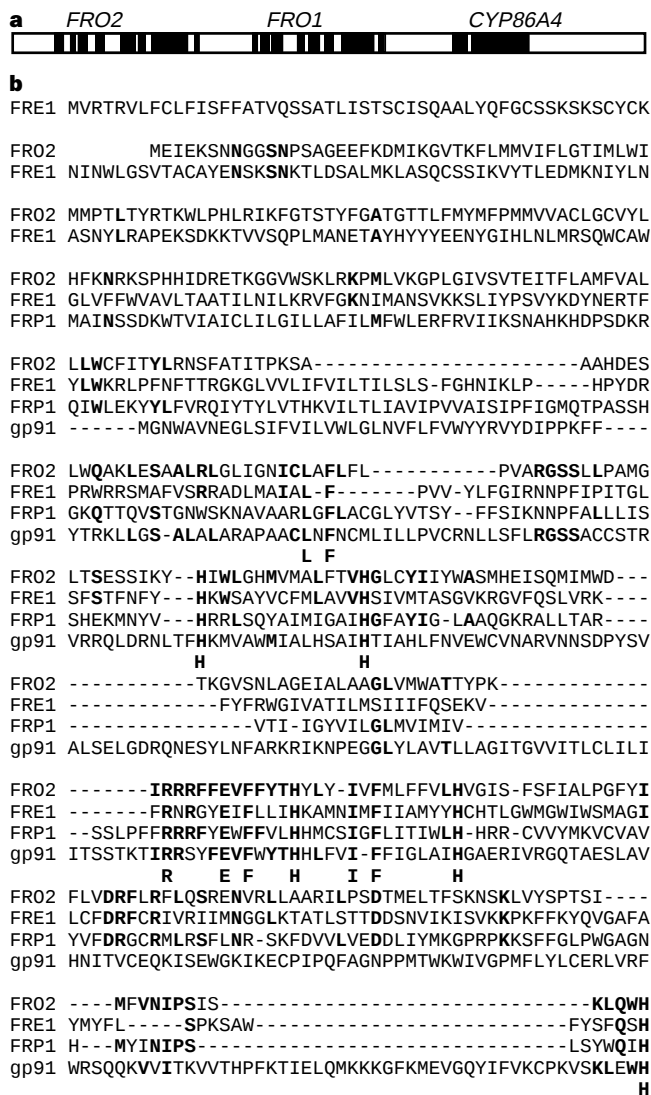
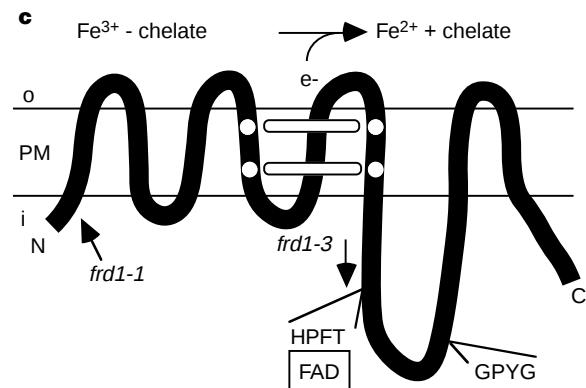


Figure 1 The sequence and predicted structure of *FRO2*. **a**, Genomic organization of *FRO2*, *FRO1* and *CYP86A4* within 15.2 kb of contiguous sequence. Black boxes represent exons. **b**, The deduced sequence of *FRO2* aligned with FRE1, FRP1 and gp91phox (gp91) with residues showing identity to *FRO2* in bold and two motifs associated with cofactor binding underlined. The *FRO2* sequence presumes



initiation at the first transcribed, in-frame start codon. **c**, Hypothetical plasma membrane (PM)-associated structure of *FRO2*. Four histidine residues (white spots) predicted to coordinate two intramembraneous haem groups (white bars) are indicated; i, inside cell; o, outside cell.

might be expected to reduce transcript stability and hence abundance. Accumulation of non-functional FRO2 in *frd1-3* may impair any feedback inhibition of transcription.

Iron uptake by leaf mesophyll cells seems to involve the action of a plasma-membrane ferric-chelate reductase, which is presumed to release the metal from ligands in the translocation stream¹⁴. Relatively little is known about the transport of iron in phloem or its uptake by developing seeds, although changes in iron demand in aerial tissues alters the activity of root-surface ferric-chelate reductase¹⁵. In addition to *FRO2* and *FRO1*, we have identified at least three other *FRO* genes in *Arabidopsis*, the products of which may act in different organs or specialized cell types to mediate the translocation of iron¹⁶. In response to iron deficiency in *Arabidopsis*, there is also increased root-surface copper-chelate reductase activity³. Low-iron-inducible copper-chelate reduction is absent in *frd1-1* mutants³ and is restored in transgenic plants containing *FRO2* (data not shown). However, *FRO2*-mediated copper reduction may be gratuitous because copper accumulation was not reduced in *frd1* mutants grown on plates³; alternatively, it may increase uptake only under certain soil conditions. Furthermore,

FRO2 transcript abundance under iron deficiency was unaffected by exposure to 0.5 μM CuSO_4 , whereas the addition of 300 μM bathocuproinedisulphonic acid (a copper chelator) further reduced, rather than increased, *FRO2* transcript abundance in plants grown under high-iron conditions for 3 days (data not shown). Some *FRE1* and *FRE2* homologues of the yeast *Saccharomyces cerevisiae* are regulated by the iron-responsive transcription factor AFT1, and some are regulated by the copper-responsive transcription factor MAC1 (ref. 17). Other *FRO* proteins may also be produced in response to low levels of copper.

We also identified an expressed sequence tag from rice with sequence similarity to gp91phox, and the corresponding gene, *rbohA*, was subsequently sequenced¹⁸. Homologues of *rbohA* have now been identified in *Arabidopsis* but their functions remain unclear^{19,20}. The deduced Rboh proteins show greater sequence identity than *FRO* to gp91phox, with sequence trees placing the *FRO* proteins in a clade separate from the Rboh proteins and gp91phox. Plants resist pathogens by means of the swift, localized production of O_2^- and H_2O_2 , which is reminiscent of the oxidative burst mediated by the gp91phox complex of human phagocytic

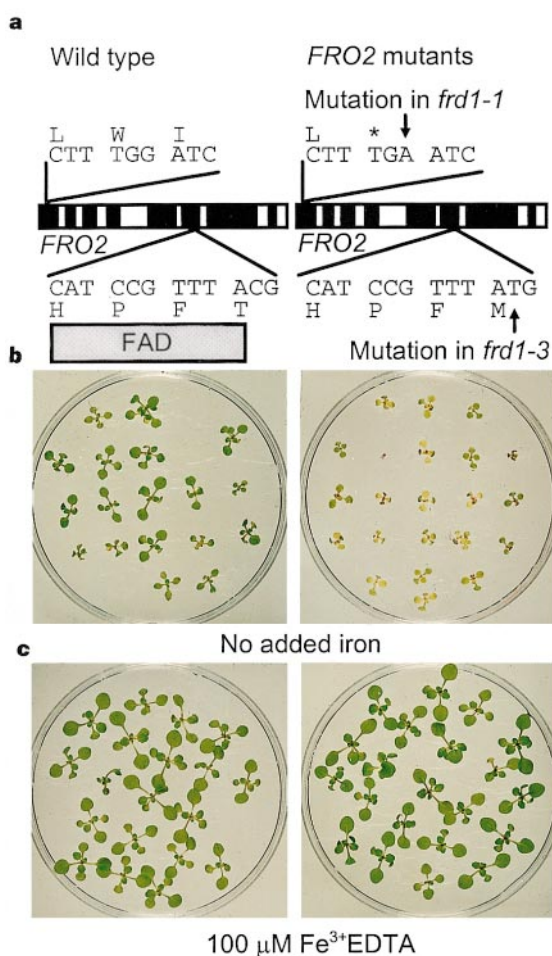


Figure 2 *FRO2* mutations in ferric-chelate reductase-deficient *frd1* *Arabidopsis* and the influence of iron on growth. **a**, Schematic representation of the *FRO2* gene (black boxes represent exons) and mutations in the first exon of *FRO2* in *frd1-1* and within the region encoding the predicted FAD-binding site in *frd1-3*. **b**, Growth of wild-type (Columbia Col-0) (left) and *frd1-1* mutants (right) for 14 days on Gamborg's agar plates containing no micronutrient iron. **c**, Growth of wild-type (Columbia Col-0) (left) and *frd1-1* (right) for 14 days on Gamborg's agar plates containing 100 μM Fe^{3+} EDTA.

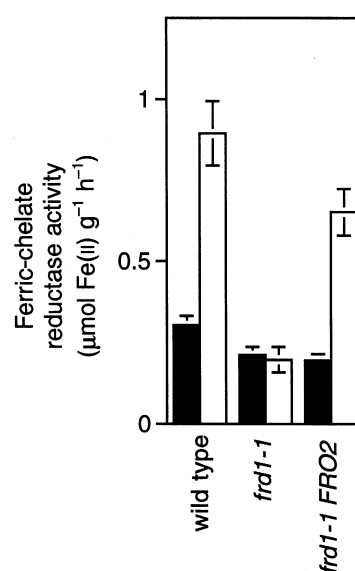


Figure 3 Assays of ferric-chelate reductase activity in wild type (Columbia *gl1*), *frd1-1* mutants and transgenic *frd1-1* mutants containing *FRO2*, showing that *FRO2* complements the *frd1* mutant phenotype. Values are mean of ten assays of ferric-chelate reductase activity in plants grown in the presence (black columns) or absence (white columns) of added Fe^{3+} EDTA using previously described conditions³.

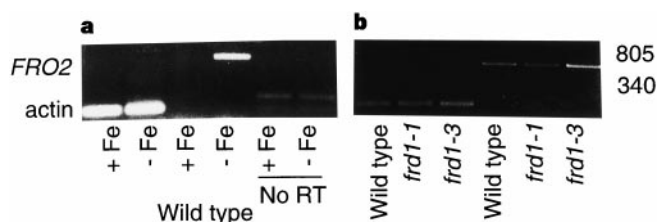


Figure 4 Semiquantitative RT-PCR showing *FRO2* transcript abundance in roots. **a**, Effects of iron (+Fe or -Fe) on transcript abundance in wild-type *Arabidopsis* (Columbia Col-0). Actin (lanes 1, 2) and *FRO2* (lanes 3, 4) amplification products are of the anticipated sizes for cDNA. A control reaction containing no reverse transcriptase (lanes 5, 6) generated a trace amount of product (faint band) of the (slightly larger) size anticipated from genomic DNA template using the actin primers. **b**, Transcript abundance under -Fe conditions in wild-type (Columbia C24) and *frd1* mutant plants using actin (lanes 1-3) and *FRO2* (lanes 4-6) primers.

neutrophil membranes. It has therefore been suggested that the products of the *rboh* genes act to generate reactive oxygen species for defence against plant pathogens^{18–20}. Data showing that at least one member (FRO2) of this superfamily of proteins is engaged in the transfer of electrons across a plant cell membrane to a single electron acceptor support the proposal that the Rboh proteins function in a similar manner but donate electrons to different acceptors.

The isolation of the *FRO* gene family has revealed a biochemical mechanism for ferric-chelate reduction by plants. We are now able to investigate the contributions of different reductase enzymes in the acquisition of exogenous iron and the distribution of iron between different cell types and organs. These genes may also be used as markers in selective breeding programmes or may be manipulated in transgenic plants to generate higher-yielding, iron-efficient or iron-rich crops. To improve human iron nutrition, it will not be sufficient merely to increase iron uptake by plants but modified uptake and/or translocation is likely to be a contributory factor. □

Methods

Gene isolation and characterization. Degenerate primers 5'-(C/T)(G/T)I (G/C)A(A/G)(A/T)II CA(C/T)CCI TT(C/T) AC-3' and 5'-IIC C(A/G)(A/T) AIG GIC CIT C-3' (where I represents deoxyinosine, and alternative deoxynucleotides at a single position are shown in parentheses) were designed to anneal to sequences encoding residues [L/F/W][Q/E][W/I/S]HPFT and [D/E]GP[Y/F]G corresponding to conserved motifs in FRE1, FRE2 (ref. 21), FRP1 and gp91phox. Genomic DNA from *Arabidopsis* (Ler) was used as template for 30 cycles of PCR with *Taq* DNA polymerase and standard buffer conditions but with 7 mM MgCl₂ and 1 μM of each primer; annealing, extension, denaturing were at 35 °C (2 min), 72 °C (1 min), 92 °C (1 min), respectively, and products were resolved on a 2% agarose gel. Seven different DNA fragments were cloned and sequenced. A probe prepared by random primer labelling²² of the insert (172 base pairs) in one of these clones (designated J1) was used to screen a mixed-tissue *Arabidopsis* cDNA library²³ by standard protocols²⁴, and one hybridizing partial cDNA (1.4 kb) was identified from 800,000 plaques. A plasmid containing the cDNA was recovered in *Escherichia coli* strain DH10B[ZIP] (Gibco BRL), the insert was sequenced and found to share 85% identity at the nucleotide level with the original PCR product, and the corresponding gene was subsequently designated *FRO3*. A probe was generated using the *FRO3* cDNA fragment and used to screen an *Arabidopsis* (Ler) genomic library²⁵ in λFIX (Stratagene). Six hybridizing clones were isolated from 200,000 plaques. Restriction mapping showed that they belonged to two groups with three clones containing the 3'-end of *FRO3*. The other three genomic clones contained various spans of another genomic region. The entire sequence of one of these genomic clones and part of another were determined to generate 15.2 kb of continuous sequence. This genomic region contains two related genes in tandem, designated *FRO2* and *FRO1*. The original PCR product, J1, was found to be identical to a section of *FRO1*. A full-length *FRO2* cDNA was subsequently obtained by screening a cDNA library prepared from RNA isolated from roots²⁶.

Semiquantitative RT-PCR. Wild-type *Columbia* (Col-0 and its variant C24) and mutant *Arabidopsis* were grown for 14 days on Gamborg's agar containing 100 μM Fe³⁺ EDTA, followed by 3 days on either no added iron but with 300 μM ferrozine (-Fe), or 100 μM Fe³⁺ EDTA (+Fe). For analysis of the effects of copper, plants were grown for the final 3 days on minimal medium agar containing no added iron either with or without 0.5 μM CuSO₄; some plants were grown for a final 6 days on minimal medium agar, with or without 300 μM bathocuproinedisulphonic acid (a copper chelator); 100 μM Fe³⁺ EDTA was added to both media for the final 3 days. Total RNA was extracted from excised roots using TRI-Reagent (Sigma), and 2 μg was used to synthesize cDNA with a dT₁₅ oligonucleotide and M-MLV reverse transcriptase (Gibco BRL). PCR with cDNA was performed using standard conditions with primers (5'-GAG ATA GAA ATC CTG AGA GG-3' and 5'-CAA AGA CCA TGA ACG GTG-3') designed to anneal to *FRO2* in preference to other *FRO* genes. Amplification of actin cDNA was used as an internal control in comparative RT-PCR experiments with primers (5'-GGT AAC ATT GTG CTC AGT GGT GG-3' and 5'-CTC GGC CTT GGA GAT CCA CAT C-3') designed to anneal to any of the ten known *Arabidopsis* actin genes²⁷.

Plant transformation and assay of ferric-chelate and copper-chelate reductase. The 6.2-kb *Xba*I genomic fragment from qnrλJ1Z4 (EMBL Y09581), containing the entire *FRO2* coding region and 0.6 kb of upstream sequence, was cloned into the *Xba*I site of pCGN1547 (ref. 28) to yield pELC203, which was used to transform *Agrobacterium tumefaciens* strain ASE to gentamycin resistance²⁹. *frd1-1* plants were grown in soil under low light conditions (50–60 μE m⁻² s⁻¹) in preparation for transformation by vacuum infiltration³⁰. Transformed plants were identified by selection on kanamycin. Ferric-chelate and copper-chelate reductase activities were assayed in wild-type *Columbia gl1, frd1-1* and the transformants, as described previously³ with the exception that plants were assayed for ferric-chelate reductase activity after four days of growth on iron-deficient or iron-sufficient media.

Received 16 October 1998; accepted 4 January 1999.

- World Health Organisation (http://www.who.int/nut/malnutrition_worldwide.htm#ida), 25 January 1999.
- Guerinot, M. L. & Yi, Y. Iron: nutritious, noxious and not readily available. *Plant Physiol.* **104**, 815–820 (1994).
- Yi, Y. & Guerinot, M. L. Genetic evidence that induction of root Fe(III) chelate reductase activity is necessary for iron uptake under iron deficiency. *Plant J.* **10**, 835–844 (1996).
- Chaney, R. L., Brown, J. C. & Tiffin, L. O. Obligatory reduction of ferric chelates in iron uptake by soybeans. *Plant Physiol.* **50**, 208–213 (1972).
- Bienfait, H. F. Regulated redox processes at the plasmalemma of plant root cells and their function in iron uptake. *J. Bioenerg. Biomembr.* **17**, 73–83 (1985).
- Eide, D. in *Metal Ions in Gene Regulation* (eds Silver, S. & Walden, W.) 342–371 (Chapman and Hall, New York, 1998).
- Chanock, S. J., Elbenna, J., Smith, R. M. & Babior, B. M. The respiratory burst oxidase. *J. Biol. Chem.* **269**, 24519–24522 (1994).
- Dancis, A., Klausner, R. D., Hinnebusch, A. G. & Barriocanal, J. G. Genetic evidence that ferric reductase is required for iron uptake in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **10**, 2294–2301 (1990).
- Roman, D. G., Dancis, A., Anderson, G. J. & Klausner, R. D. The fission yeast ferric reductase gene *frp1* is required for ferric iron uptake and is homologous to the gp91-phox subunit of the human NADPH phagocyte oxidoreductase. *Mol. Cell. Biol.* **13**, 4342–4350 (1993).
- Finegold, A. A., Shattwell, K. P., Segal, A. W., Klausner, R. D. & Dancis, A. Intramembrane bis-heme motif for trans-membrane electron transport conserved in a yeast iron reductase and the human NADPH oxidase. *J. Biol. Chem.* **271**, 31021–31024 (1996).
- Solovye, V. V., Salamov, A. A. & Lawrence, C. B. Predicting internal exons by oligonucleotide composition and discriminant analysis of spliceable open reading frames. *Nucleic Acids Res.* **22**, 5156–5163 (1994).
- Milpetz, F., Argos, P. & Persson, B. TMAP- a new email and www service for membrane-protein structural predictions. *Trends Biochem. Sci.* **20**, 204–205 (1995).
- Arabidopsis thaliana* Genome Center (http://genome.bio.upenn.edu/physical-mapping/ch1_pt1.html), 25 January 1999.
- Brüggenman, W., Mass-Kantel, K. & Moog, P. R. Iron uptake in mesophyll cells: the role of plasma membrane bound ferric-chelate reductases. *Planta* **190**, 151–155 (1993).
- Grusak, M. A. & Pezeshgi, S. Shoot-to-root signal transmission regulates root Fe(III) reductase activity in the *dgl* mutant of pea. *Plant Physiol.* **110**, 329–334 (1996).
- Robinson, N. J., Sadjuga & Groom, Q. J. The *frh* gene family from *Arabidopsis thaliana*: putative iron-chelate reductases. *Plant Soil* **196**, 245–248 (1997).
- Martins, L. J., Jensen, L. T., Simons, J. R., Keller, G. L. & Winge, D. R. Metalloregulation of *FRE1* and *FRE2* homologs in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **273**, 23716–23721 (1998).
- Groom, Q. J. *et al.* *rbohA*, a rice homologue of the mammalian gp91-phox respiratory burst oxidase gene. *Plant J.* **10**, 515–522 (1996).
- Keller, T. *et al.* A plant homolog of the neutrophil NADPH oxidase gp91phox subunit gene encodes a plasma membrane protein with Ca²⁺ binding motifs. *Plant Cell* **10**, 1–13 (1998).
- Torres, M. A. *et al.* Six *Arabidopsis thaliana* homologues of the human respiratory burst oxidase (gp91phox). *Plant J.* **14**, 365–370 (1998).
- Georgatsou, E. & Alexandraki, D. Two distinctly regulated genes are required for ferric reduction, the first step of iron uptake in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **14**, 3065–3073 (1994).
- Feinberg, A. P. & Vogelstein, B. A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **132**, 6–13 (1983).
- Newman, T. *et al.* Genes galore: a summary of the methods for accessing results from large-scale partial sequencing of anonymous *Arabidopsis* cDNA clones. *Plant Physiol.* **106**, 1241–1255 (1994).
- Benton, W. D. & Davis, R. W. Screening Agt recombinant clones by hybridisation to single plaques *in situ*. *Science* **196**, 180–182 (1977).
- Voytas, D. E., Konieczny, A., Cummings, M. P. & Ausubel, F. M. The structure, distribution and evolution of the *Ta1* retrotransposable element family of *Arabidopsis thaliana*. *Genetics* **126**, 713–721 (1990).
- Peterman, T. K. & Goodman, H. M. The glutamine synthetase gene family of *Arabidopsis thaliana*: light-regulation and differential expression in leaves, roots and seeds. *Mol. Gen. Genet.* **230**, 145–154 (1991).
- McDowell, J. M., Huang, S. R., McKinney, E. C., An, Y. Q. & Meagher, R. B. Structure and evolution of the actin gene family in *Arabidopsis thaliana*. *Genetics* **142**, 587–603 (1996).
- McBride, K. E. & Summerfelt, K. R. Improved binary vectors for *Agrobacterium*-mediated plant transformation. *Plant Mol. Biol.* **14**, 269–276 (1990).
- Rogers, S. G., Klee, H. J., Horsch, R. B. & Fraley, R. T. Improved vectors for plant transformation: expression cassette vectors and new selectable markers. *Methods Enzymol.* **253**, 253–277 (1988).
- Bent, A. F. *et al.* *RPS2* of *Arabidopsis thaliana*: a leucine-rich repeat class of plant disease resistance genes. *Science* **265**, 1856–1860 (1994).

Acknowledgements. We thank Q. J. Groom, funded by BBSRC (N.J.R.), whose work provided a basis for this study. This work was supported by a BBSRC studentship (C.M.P.) and a grant from the US National Science Foundation (M.L.G.).

Correspondence and requests for materials should be addressed to N.J.R. (requests for *FRO* clones; e-mail: n.j.robinson@ncl.ac.uk) or M.L.G. (*frd1* mutants; e-mail: mary.lou.guerinot@dartmouth.edu). *FRO1* and *FRO2* sequence has been deposited in the EMBL database (accession number Y09851).

A diffusion barrier maintains distribution of membrane proteins in polarized neurons

Bettina Winckler*, Paul Forscher† & Ira Mellman*

* Department of Cell Biology, Yale University School of Medicine, 333 Cedar Street, PO Box 208002, and † MCDB, Yale University, New Haven, Connecticut 06520-8002, USA

The asymmetric distribution of proteins to distinct domains in the plasma membrane is crucial to the function of many polarized cells. In epithelia, distinct apical and basolateral surfaces are maintained by tight junctions that prevent diffusion of proteins and lipids between the two domains¹. Polarized neurons maintain axonal and somatodendritic plasma membrane domains without an obvious physical barrier. Indeed, the artificial lipid DiI encounters no diffusion barrier at the presumptive domain boundary, the axon hillock². By measuring the lateral mobility of membrane proteins using optical tweezers, we show here that some membrane proteins exhibit markedly reduced mobility in the initial segment of the axon. Disruption of F-actin and low levels of dimethyl sulphoxide (DMSO) abolish this diffusion barrier and lead to redistribution of membrane markers that had previously been polarized. Immobilization in the initial segment may reflect, at least in part, differential tethering to cytoskeletal components. Therefore, the ability to maintain a polarized distribution of membrane proteins depends on a specialized domain at the initial segment of the axon, which restricts lateral mobility and serves as a new type of diffusion barrier that acts in the absence of cell–cell contact.

Little is known about how the asymmetric distribution of axonal and somatodendritic proteins is maintained. After polarized delivery to the axon, axonal plasma membrane proteins might be retained by means of cytoskeletal tethering. Alternatively, proteins might be freely diffusible in the plasma membrane, but quickly retrieved from the incorrect domain by endocytosis. Finally, proteins might be freely diffusible, but encounter a diffusion barrier of a kind analogous to epithelial tight junctions at the boundary between the axonal and somatodendritic compartments.

To map the boundary between axonal and somatodendritic compartments, sparse cultures of rat hippocampal neurons were immunostained for the axonal cell-adhesion molecule L1 and the somatodendritic glutamate receptor subunit GluR1 (Fig. 1a, b). The staining against GluR1 (red) extended $33 \pm 20 \mu\text{m}$ (mean $\pm$ standard deviation) beyond the axon hillock (outlined in Fig. 1b),

well into the proximal region of the axon. Conversely, L1 (green) ceased to be detectable $30 \mu\text{m}$ ($\pm 15 \mu\text{m}$) distal to the soma boundary. Similar results were obtained for other axonal markers including NgCAM, an avian L1 homologue expressed using a recombinant adenovirus (AdNgCAM³), and for the endogenous glycosyl phosphatidylinositol (GPI)-anchored protein Thy1 (data not shown). Thy1 is axonal in only 30–40% of neurons (but see ref. 4). These observations indicate that the axon initial segment, rather than the axon hillock⁵, comprises the boundary between axon and soma, possibly by impeding the diffusion of membrane proteins between the two domains.

To test for a diffusion barrier at the initial segment, we used optical tweezers which have been used successfully in other cells to study the lateral mobility of membrane proteins⁶. We placed antibody-coupled beads with optical tweezers on the axonal process at various distances from the soma. After attachment, the beads were trapped again and dragged along the axon. Fig. 2a shows the tractability of anti-Thy1 beads on a 20-day-old neuron. The tractability of beads correlated well with their position on the axon: beads placed on the initial segment could on average be dragged $0.28 \mu\text{m}$, whereas beads placed further out on the axon (more than $70\text{--}90 \mu\text{m}$ from the soma) could be dragged on average of $3.7 \mu\text{m}$. Fig. 2b summarizes the results for three markers: the mobilities of a transmembrane protein (L1) and a GPI-linked protein (Thy1) were reduced to the same degree on the initial segment compared with the main axon. Similar results were obtained using concanavalin A (ConA)-coated beads, raising the possibility that many other membrane proteins display reduced lateral mobility at the initial segment. ConA-coupled beads and anti-Thy1 beads placed on dendrites were mostly tractable ($68 \pm 16\%$ for ConA; $94 \pm 13\%$ for Thy1), independent of distance.

To determine whether F-actin is important in restricting lateral mobility in the initial segment, we used the actin-disrupting drug, latrunculin-B (LAT-B). In control experiments, however, we found that the solvent DMSO alone showed an effect. The barrier at the initial segment that is encountered by both L1 and Thy1 is destroyed by 0.4% DMSO (Fig. 2c). DMSO had no obvious effects on cell morphology, actin-based growth cone motility or axonal organelle transport as recorded by video-enhanced contrast/differential interference contrast (DIC) microscopy. We did observe that DMSO facilitated the ability to pull thin membrane tethers up to $15 \mu\text{m}$ in length with the optical tweezers (in $\sim 50\%$ of the L1 beads). These tethers could be pulled both on the initial segment and the main axon, but not in untreated control cells. It has been reported previously that low levels of DMSO lowered the force needed for tether formation⁷. Microtubules, F-actin and neurofilaments all appeared normal by immunofluorescence in DMSO-treated cultures (see Supplementary Information). We surmise that DMSO uncouples the membrane from the underlying membrane skeleton, allowing both tether formation and increased bead tractability.

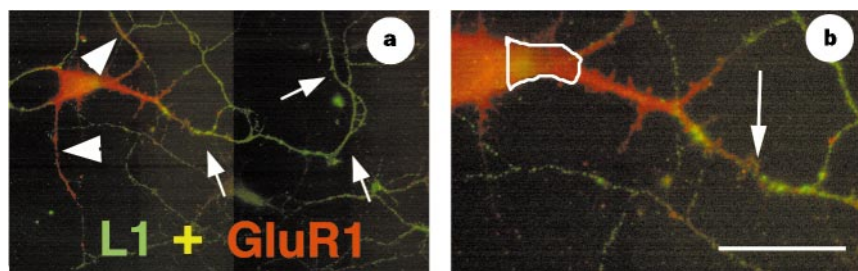


Figure 1 The axonal initial segment compartmentalizes axonal and somatodendritic domains. **a**, Hippocampal neurons (after 9–11 days in culture) were double-stained with antibodies to the endogenous markers L1 (green) and GluR1 (red). **b**, Magnification of the axon hillock and region of the initial segment of the cell

shown in **a**; the axon hillock is outlined. Small arrows indicate axons; arrowheads mark dendrites and the large arrow (in **b**) indicates the initial segment (defined operationally as the initial $\sim 80 \mu\text{m}$ of axon minus the tapering axon hillock). Scale bars, $20 \mu\text{m}$.

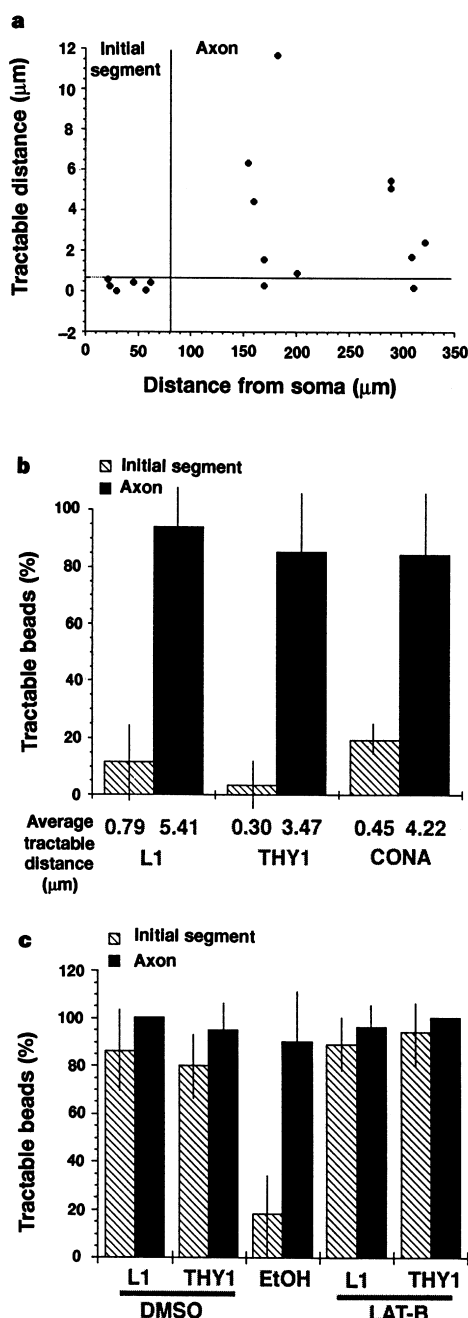


Figure 2 Bead mobility is restricted in the axon initial segment in control but not DMSO- or LAT-treated cells. **a**, Bead tractability of 17 anti-Thy1 beads placed at various distances from the soma of a single 20-day-old neuron. The distance each bead was dragged using an optical tweezer (tractable distance) was plotted against its position on the axon (distance from soma). The total axon length of this cell was $>400\mu\text{m}$. Whereas 82% of beads placed on the distal axon were tractable, none was tractable in the initial segment. **b**, Bead probes to three membrane proteins display reduced tractability in the initial segment. The percentage of beads that were immobile was measured for the initial segment (striped bars) and axon (solid bars) of each cell and the mean and standard deviation of several experiments were plotted for anti-L1 beads (eight cells), anti-Thy1 beads (seven cells) and ConA beads (four cells). Average tractable distances are indicated beneath each bar. **c**, DMSO (0.4%) and $0.4\mu\text{g ml}^{-1}$ LAT-B increased the tractability of both anti-Thy1 beads and anti-L1 beads in the initial segment. The means and standard deviation were plotted for anti-L1 beads (DMSO, four cells; LAT-B, four cells) and anti-Thy1 beads (DMSO, five cells; LAT-B, five cells). Ethanol controls were identical to controls without ethanol (five cells). IS = striped bars; axon = solid bars.

We used ethanol as a solvent for LAT-B. F-actin staining was greatly disrupted in LAT-B. Similarly to DMSO, cells treated with $0.4\mu\text{g ml}^{-1}$ LAT-B for 1.5–3 h showed no region of reduced tractability at the initial segment for L1 and Thy1 (Fig. 2c). Cells treated with the same concentration of ethanol alone were identical to controls without ethanol (Fig. 2c). The barrier recovered fully for L1 when LAT-B was washed out for several hours. For Thy1, the reversal was variable (not shown).

We tested the physiological significance of the diffusion barrier of the initial segment: in the presence of DMSO or LAT-B, neurons should not be able to maintain the polarized distribution of plasma membrane components. We infected neurons with AdNgCAM³. Treatment of infected cultures with 0.4% DMSO or $0.4\mu\text{g ml}^{-1}$ LAT-B for 24 h resulted in extensive randomization of NgCAM distribution (Fig. 3a, b), even in the presence of cycloheximide (Fig. 3c). The appearance of NgCAM on the somatodendritic plasma membrane in the presence of DMSO or LAT-B was, thus, due to the relocalization of pre-existing molecules. The loss of polarity of NgCAM in DMSO takes several hours and increases with time (see Supplementary Information), which is consistent with diffusion. Similarly, the somatodendritic protein GluR1 was less well polarized in DMSO (see Supplementary Information). In contrast, the localization of microtubule-associated protein 2 (MAP2)—a somatodendritic cytoskeletal protein—was not affected (not shown). Other actin-dependent processes (such as sorting in endosomes⁸), which are potentially important for the polarized distribution of L1, might in addition contribute to the observed loss of polarity. These results indicate that the axonal initial segment acts as a DMSO- and LAT-sensitive diffusion barrier and maintains the polarized distribution of axonal and somatodendritic membrane proteins.

Indeed, the initial segment seems to act by physically restricting the diffusion of L1 trapped within it. To quantify diffusion, we performed high resolution analysis of bead trajectories by single-particle tracking⁶, revealing a 50-fold difference in microscopic diffusion coefficients (D_{2-4}) (as defined by ref. 9) between immobile and tractable L1 beads (average immobile beads $D_{2-4} = 7.6 \times 10^{-12} \pm 9.6\text{ cm}^2\text{ s}^{-1}$; average tractable beads $D_{2-4} = 1.3 \times 10^{-9} \pm 0.85\text{ cm}^2\text{ s}^{-1}$). A total of 80% of L1 beads on the IS showed slow diffusional mobility ($D_{2-4} < 5 \times 10^{-11}\text{ cm}^2\text{ s}^{-1}$). In every case, these were the non-tractable beads using the optical tweezer assay. All beads analysed on the axon showed 20–100-fold-higher diffusion coefficients ($D_{2-4} = 8.9 \times 10^{-10} \pm 5.6\text{ cm}^2\text{ s}^{-1}$). In the presence of DMSO, all beads showed high diffusional mobility regardless of position ($D_{2-4} = 1.6 \times 10^{-9} \pm 1.0\text{ cm}^2\text{ s}^{-1}$). Thus, most L1 molecules in the initial segment diffused slowly or not at all. Because some ($\sim 20\%$) of the L1 beads were tractable in the initial segment and exhibited high diffusional mobility, a subset of L1 molecules diffused rapidly, possibly accounting for the steady-state residency of L1 in the initial segment. By removing the restriction on lateral mobility, DMSO allows free diffusion through the initial segment leading to the observed randomized distribution of polarized membrane proteins.

The restricted diffusion of membrane proteins in the initial segment suggests direct or indirect tethering to the cytoskeleton. This is further supported by the observation that intact F-actin is required for barrier function. As tethered proteins are often insoluble in non-ionic detergents, neurons expressing NgCAM were extracted in 1% CHAPS before fixation and immunostaining. Although NgCAM staining at the initial segment was normally less bright than in the rest of the axon, after CHAPS extraction the brightest NgCAM staining was at the initial segment (Fig. 4a, b; arrows). Similar results were obtained for endogenous L1 after CHAPS extraction at 4°C and 37°C . The distal extent of the CHAPS-resistant segment ($88 \pm 25\mu\text{m}$ for L1; $78 \pm 10\mu\text{m}$ for NgCAM) was similar to the length of the restricted diffusion zone measured with the optical tweezers ($60\text{--}90\mu\text{m}$). These results

suggest that L1 interacts differently with the cytoskeleton on the initial segment and in the rest of the axon. When the cells were treated with DMSO or LAT-B before CHAPS extraction, L1 became more fully extractable in the initial segment (Fig. 4c). For LAT, this effect was less complete than for DMSO, but we observed that the staining intensity of L1 in the initial segment was reduced in LAT-treated cells scored as initial-segment-enriched (see Supplementary Information) compared with controls. Thy1, however, was CHAPS-resistant to the same extent in the initial segment and axon at 4 °C, but was fully extracted at 37 °C (not shown). This behaviour is consistent with the observation that GPI-anchored proteins partition into lipid rafts¹⁰ and are unable to link directly to the cytoskeleton.

Electron micrographs of brain sections have revealed that the initial segment, like nodes of Ranvier, contains a specialized membrane cytoskeleton, which can be seen as a tripartite undercoating

beneath the plasma membrane¹¹. In cultured neurons, this specialization is not apparent¹². Two possible components of the undercoating have been identified: a neuronal splice isoform of ankyrin_G (relative molecular mass, 480/270K)¹³ and an isoform of amphiphysin II (ref. 14). In cultured hippocampal neurons, ankyrin_{G480/270} was highly enriched at the initial segment, whereas amphiphysin II was somewhat enriched. After CHAPS extraction, however, both amphiphysin II and ankyrin_{G480/270} remained behind and largely co-localized in the initial segment (see Supplementary Information). F-actin likewise remained in the initial segment after CHAPS extraction (see Supplementary Information). Thus, these two proteins in conjunction with F-actin may participate in the detergent-insoluble complex that seems to restrict L1 diffusion in the initial segment. In contrast to L1, though, their solubility is not affected by previous DMSO treatment (not shown). Ankyrin is known to interact with the cytoplasmic domain of L1-family

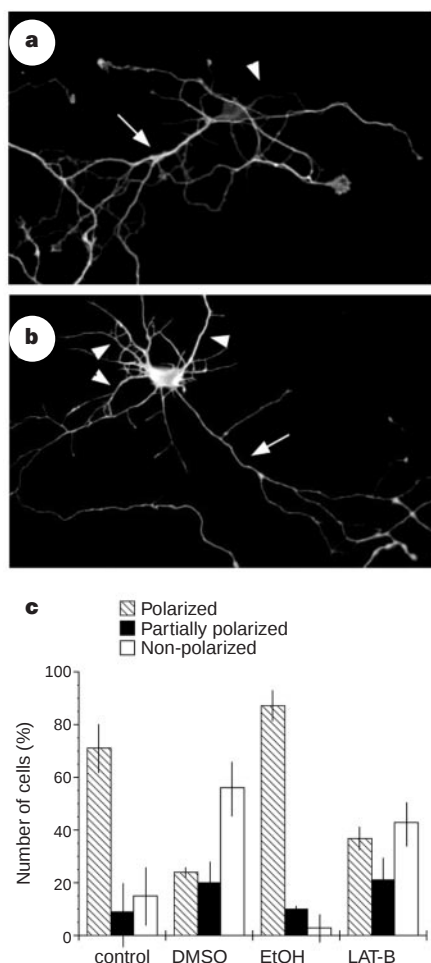


Figure 3 DMSO and LAT-B treatment lead to randomization of axonal NgCAM. **a**, **b**, Ten-day-old neuronal cultures were infected for 2 days with AdNgCAM, treated with (+) or without (-) DMSO for 24 h and then processed for immunofluorescence using an anti-NgCAM antibody. **a**, In controls, NgCAM largely remains axonal (arrows; defined as 'polarized' in **c**). The background staining on the soma in this panel is due to nonspecific binding of the antibody and is comparable to the staining of uninfected cells. **b**, In DMSO-treated cells, NgCAM staining was also found on the soma and dendrites (arrowheads) ('nonpolarized' in **c**). **c**, NgCAM distribution was quantified in control, DMSO- and LAT-treated cells for three experiments each (average $\pm$ s.d.). Ten-day-old neurons were infected as above but then treated with 20 μ g ml⁻¹ cycloheximide and $\pm$ 0.4% DMSO, 1:50,000 ethanol or 0.4 μ g ml⁻¹ LAT-B for 24 h before fixation. Cells were scored as 'polarized' (no somatodendritic staining; striped bars), 'partially polarized' (some somatodendritic staining; black bars) or 'nonpolarized' (somatodendritic staining intensity comparable to axonal staining; white bars).

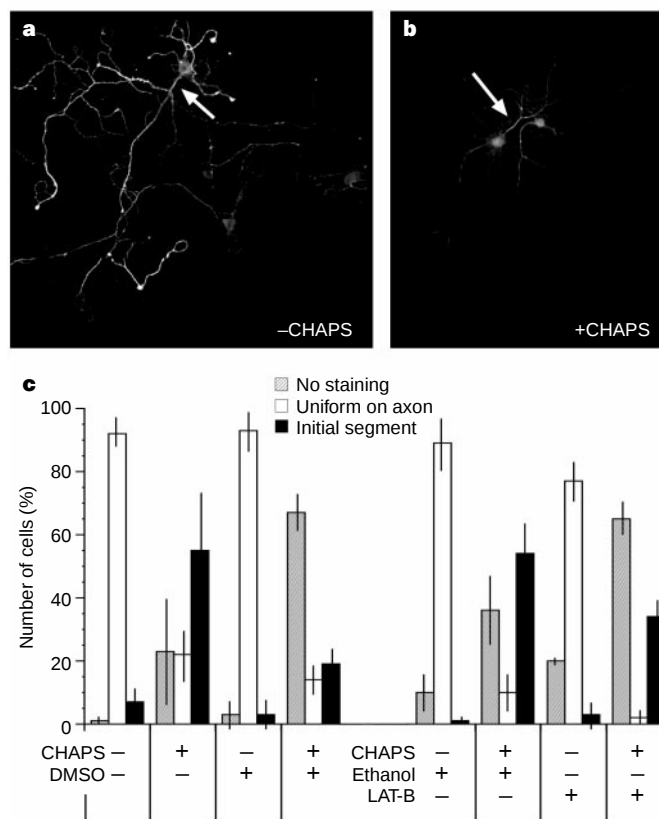


Figure 4 L1 and NgCAM are part of a DMSO- and LAT-sensitive CHAPS-resistant complex at the initial segment. **a**, AdNgCAM-infected, day-11 neurons were fixed without previous CHAPS treatment and stained for NgCAM. NgCAM was found throughout the axon, extending to the distal tips. The initial segment is indicated by an arrow. **b**, Infected day-11 neurons were fixed after treatment with 1% CHAPS and stained for NgCAM. NgCAM was largely extracted from distal axonal regions but retained in the initial segment (arrow). **c**, Quantification of results obtained in three separate experiments for the distribution of anti-L1 staining with or without CHAPS extraction, with or without 1-h incubation in 0.4% DMSO, with 1:50,000 ethanol or with 0.4 μ g ml⁻¹ LAT-B for 3 h (average $\pm$ s.d.). Staining was scored as axonal (empty bars; for example, cell in **a**), initial segment (black bars; for example, cell in **b**) or no staining (striped bars, staining not visible).

members¹⁵. Moreover, the ankyrin interaction of the L1-family member neurofascin is regulated by phosphorylation at a conserved cytoplasmic tail tyrosine residue¹⁶. Mutation of this site leads to increased lateral mobility of neurofascin in neuroblastoma cells.

Therefore, we propose that the specialized spectrin/actin/ankyrin skeleton in the initial segment not only anchors membrane proteins such as neurofascin and sodium channels¹⁷, but also creates a diffusion barrier for other membrane proteins. L1 seems freely diffusible in the main axon, but becomes tethered when entering the initial segment. Specific endocytic retrieval in the axonal initial segment or elsewhere might additionally be important for restricting the distribution of L1 to the axon. How Thy1 diffusion is restricted is less clear, as it does not associate with the CHAPS-resistant cytoskeleton at 37 °C. Thy1 diffusion might be impeded by the great density of proteins (that is, neurofascin¹⁷ and sodium channels^{17–19}) directly tethered in the initial segment²⁰. Because the diffusion barrier develops in the absence of myelination or glial cell interaction, its formation must be an intrinsic property of neurons. Analogous diffusion barriers might act to restrict diffusion in other cases where components are segregated in continuous membranes, for example smooth versus rough endoplasmic reticulum or anterior versus posterior tail domains in mammalian sperm²¹. □

Methods

Cell culture. Primary cultures from E18 rat embryos were prepared as previously described²² with modification²³. Cells were used after 8–14 days in culture. For experiments using optical tweezers, cells were cultured on photo-etched coverslips (Bellco Glass). Those cells were chosen where the axon appeared to emerge directly from the soma rather than from a dendrite¹². The axon hillock was identified morphologically as the tapering part of the axon before it becomes cylindrical.

Reagents. The following antibodies were obtained: anti-MAP2 (clone HM2) from Sigma, anti-GluR1 (AB1504) from Chemicon, anti-L1 (ASC54) and anti-NgCAM (8D9) from the NIH Hybridoma Bank, anti-Thy1 (OX-7) and anti-tubulin (YOL 1/34) from PharMingen, and anti NF-M (13-0800) from Zymed. Antibodies against ankyrin_{G480/270} (ref. 13) and amphiphysin II (ref. 14) and the recombinant adeno-NgCAM³ were provided by V. Bennett, P. de Camilli and P. Sonderegger, respectively. LAT-B was obtained from Calbiochem.

Bead preparation. Amino-modified silica beads (Bangs Laboratories) of size 0.5 µm were coupled to protein G as described²⁴. Antibody-coupled beads were prepared fresh for each experiment. Between 4,000 and 10,000 antibody molecules were coupled to each bead. Beads coupled to anti-Thy1, anti-L1 or conA bound well to axons (60–90% of beads bound), whereas control beads (coupled to horse serum or protein G) bound poorly (9–15% bound). Similarly, anti-L1 beads did not bind well to dendrites (18% bound) or to L1-negative NRK cells.

Optical-tweezer experiments. Coverslips were mounted in live-cell chambers²⁵ in mammalian Ringer (140 mM NaCl, 5 mM KCl, 1 mM CaCl₂, 1 mM MgCl₂, 24 mM glucose, 10 mM Hepes, pH 7.4), sealed with vacuum grease and maintained at 36 °C. Video-enhanced DIC and image processing were performed as described previously²⁵. Cells with unambiguous axon/dendrite morphology were chosen under the 20× objective, and their position recorded. An infrared laser trap was used in these studies and was similar to that described previously²⁵. Tractability of beads was probed within 30 s of stable attachment as clustering of CAMs by beads can lead to attachment to cytoskeleton over time²⁵. Bead dissociation was observed only occasionally. For measuring tractability, the trap was moved manually at a speed of ~1 µm s⁻¹ towards the bead and past it. The distance of bead movement in the trap was determined from video-tape records using software written by P.F. A tractable bead was defined as one which could be dragged >0.6 µm. For most cells, the tractability of 10–20 beads was probed at various distances from the soma. Actual drag distances depend in many cases on physical features of the cell (for example, crossing neurites and turns). After the experiment, cells were fixed with 4% paraformaldehyde/3% sucrose/PBS and stained for MAP2 to confirm the correct identification of the axonal process. For DMSO experiments, 0.4% DMSO was added to the medium at the time of chamber construction (~1 h in DMSO). For LAT-B experiments, 0.4 µg ml⁻¹ LAT-B (stock 20 mg ml⁻¹ in

ethanol) was added for 1–3 h before the experiment. Control chambers for the LAT-B experiments contained 1:50,000 ethanol.

Single-particle tracking. Single-particle tracking was done essentially as described in ref. 9. Consecutive video frames (350 or 950; 10 or 30 s respectively) of freely diffusing beads were recorded onto optical magnetic disc recorder and the centre of the bead was determined by a centroid program written by P.F. Mean square displacements (MSD) were calculated for intervals 2–90 (0.066–3 s). D_{2-4} is one quarter of the slope of a line fit through the first three points of the MSD plot. The limit of resolution is 5×10^{-13} cm² s⁻¹ for an immobile object on the coverslip.

Other methods. For delocalization experiments, cells were infected with a recombinant adenovirus in the presence of 1 µM AraC and 0.5 mM kynurenic acid²⁶. The infections were done at a low multiplicity of infection (100 pfu per cell) so that the axons of individual infected cells could be identified unambiguously. After 48 h, the medium was replaced with 20 µg ml⁻¹ cycloheximide containing, in addition, 0.4% DMSO, 1:50,000 ethanol or 0.4 µg ml⁻¹ LAT-B for 24 h. Fresh DMSO was added after 8–12 h. Cells were fixed as above and stained without permeabilization. Extractions were done in 1% CHAPS or 1% Triton-X100 in cytoskeletal buffer (2 mM MgCl₂, 10 mM EGTA, 60 mM Pipes, pH 7.0) or PBS for 2–5 min at 4 °C or 37 °C.

Received 14 October; accepted 14 December 1998.

1. Mellman, I. Molecular sorting of membrane proteins in polarized and non-polarized cells. *Cold Spring Harb. Symp. Quant. Biol.* **60**, 745–752 (1996).
2. Winckler, B. & Poo, M.-m. No diffusion barrier at axon hillock. *Nature* **379**, 213 (1996).
3. Vogt, L. et al. Continuous renewal of the axonal pathway sensor apparatus by insertion of new sensor molecules into the growth cone membrane. *Curr. Biol.* **6**, 1153–1158 (1996).
4. Dotti, C. G., Parton, R. G. & Simons, K. Polarized sorting of glypiated proteins in hippocampal neurons. *Nature* **349**, 156–161 (1991).
5. Kobayashi, T., Storrer, B., Simons, K. & Dotti, C. G. A functional barrier to movements of lipids in polarized neurons. *Nature* **359**, 647–650 (1992).
6. Kusumi, A., Sako, Y., Fujiwara, T. & Tomishige, M. Application of laser tweezers to studies of the fences and tethers of the membrane skeleton that regulate the movements of plasma membrane proteins. *Methods Cell Biol.* **55**, 173–194 (1998).
7. Dai, J. & Sheetz, M. P. Mechanical properties of neuronal growth cone membranes studied by tether formation with laser optical tweezers. *Biophys. J.* **68**, 988–996 (1995).
8. Lamaze, C., Fujimoto, L. M., Yin, H. L. & Schmid, S. L. The actin cytoskeleton is required for receptor-mediated endocytosis in mammalian cells. *J. Biol. Chem.* **272**, 20332 (1997).
9. Kusumi, A., Sako, Y. & Yamamoto, M. Confined lateral diffusion of membrane receptors as studied by single particle tracking (nanovid microscopy). Effects of calcium-induced differentiation in cultured epithelial cells. *Biophys. J.* **65**, 2021–2040 (1993).
10. Simons, K. & Ikonen, E. functional rafts in cell membranes. *Nature* **387**, 569–572 (1997).
11. Peters, A., Proskauer, C. C. & Kaiserman-Abramof, I. R. The small pyramidal neuron of the rat cerebral cortex: the axon hillock and initial segment. *J. Cell Biol.* **39**, 604–619 (1968).
12. Bartlett, W. P. & Banker, G. A. An electron microscopic study of the development of axons and dendrites by hippocampal neurons in culture. I. Cells which develop without intercellular contacts. *J. Neurosci.* **4**, 1944 (1984).
13. Kordeli, E., Lambert, S. & Bennett, V. Ankyrin_G: a new ankyrin gene with neural-specific isoforms localized at the initial segment and Node of Ranvier. *J. Biol. Chem.* **270**, 2352–2359 (1995).
14. Butler, M. H. et al. Amphiphysin II (SH3P9/BIN1), a member of the amphiphysin/Rvs family, is concentrated in the cortical cytomatrix of axon initial segments and Nodes of Ranvier in brain and around T tubules in skeletal muscle. *J. Cell Biol.* **137**, 1355–1367 (1997).
15. Davis, J. Q. & Bennett, V. Ankyrin binding activity shared by neural fasciclin/L1/NrCAM family of nervous system cell adhesion molecules abolishes ankyrin binding and increases lateral mobility of neurofascin. *J. Cell Biol.* **137**, 703–714 (1997).
16. Garver, T. D., Ren, Q., Tuvia, S. & Bennett, V. Tyrosine phosphorylation at a site highly conserved in the L1 family of adhesion molecules abolishes ankyrin binding and increases lateral mobility of neurofascin. *J. Cell Biol.* **137**, 703–714 (1997).
17. Zhou, D. et al. Ankyrin G is required for clustering of voltage-gated Na channels at axon initial segments and for normal action potential firing. *J. Cell Biol.* **143**, 1295 (1998).
18. Angelides, K. J., Elmer, L. W., Loftus, D. & Elson, E. Distribution and lateral mobility of voltage-dependent sodium channels in neurons. *J. Cell Biol.* **106**, 1911 (1988).
19. Dargent, B. et al. Targeting of the voltage-dependent sodium channel to the axon of cultured hippocampal neurons. *Soc. Neurosci. Abstr.* **24**, 1078 (1998).
20. Sheetz, M. P. Glycoprotein mobility and dynamic domains in fluid membranes. *Annu. Rev. Biophys. Biomol. Struct.* **22**, 417 (1993).
21. Bartles, J. R. The spermatid membrane comes of age. *Trends Cell Biol.* **5**, 400–407 (1995).
22. Banker, G. & Goslin, K. (eds) *Culturing Nerve Cells* (MIT Press, Cambridge, Massachusetts, 1997).
23. Brewer, G. J., Torricelli, J. R., Evege, E. K. & Price, P. J. Optimized survival of hippocampal neurons in B-27 supplemented Neurobasal, a new serum-free medium combination. *J. Neurosci. Res.* **35**, 567–576 (1993).
24. Thompson, C., Lin, C. H. & Forscher, P. An *Aplysia* cell adhesion molecule associated with site-directed actin filament assembly in neuronal growth cones. *J. Cell Sci.* **109**, 2843–2854 (1996).
25. Forscher, P. & Smith, S. J. Actions of cytochalasins on the organization of actin filaments and microtubules in a neuronal growth cone. *J. Cell Biol.* **107**, 1505–1516 (1988).
26. Jareb, M. & Banker, G. The polarized sorting of membrane proteins expressed in cultured hippocampal neurons using viral vectors. *Neuron* **20**, 855 (1998).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements. We thank P. de Camilli, P. Sonderegger and V. Bennett for sharing their reagents; P. de Camilli, D. Suter and members of I.M.'s laboratory for advice and stimulating discussions; and F. Solomon, S. Cash, M. Velleca, H. Fölsch and R. Kroschewski for critical reading of the manuscript.

Correspondence and requests for materials should be addressed to I.M. (e-mail: ira.mellman@yale.edu).

Development of peripheral lymphoid organs and natural killer cells depends on the helix-loop-helix inhibitor Id2

Yoshifumi Yokota[†], Ahmed Mansouri^{*}, Seiichi Mori[†], Seiichi Sugawara[†], Satoko Adachi[†], Shin-ichi Nishikawa[†] & Peter Gruss^{*}

^{*} Department of Molecular Cell Biology, Max-Planck Institute of Biophysical Chemistry, Am Fassberg, 37077 Göttingen, Germany

[†] Department of Molecular Genetics, Graduate School of Medicine, Kyoto University, Shogoin Kawahara-cho 53, Sakyo-ku, 606-8507 Kyoto, Japan

Transcription factors with a basic helix-loop-helix (HLH) motif have been shown to be crucial for various cell differentiation processes during development of multicellular organisms¹. Id proteins inhibit the functions of these transcription factors in a dominant-negative manner by suppressing their heterodimerization partners through the HLH domains²⁻⁴. Members of the Id family also promote cell proliferation^{4,5}, implying a role in the control of cell differentiation. Here we show that Id2 is indispensable for normal development of mice. *Id2*^{-/-} mice lack lymph nodes and Peyer's patches. However, their splenic architecture is normal, exhibiting T-cell and B-cell compartments and distinct germinal centres. The cell population that produces lymphotoxins, essential factors for the development of secondary lymphoid organs⁶⁻¹¹, is barely detectable in the *Id2*^{-/-} intestine. Furthermore, the null mutants show a greatly reduced population of natural killer (NK) cells, which is due to an intrinsic defect in NK-cell precursors. Our results indicate that Id2 has an essential role in the generation of peripheral lymphoid organs and NK cells.

The *Id2*-gene³ locus was inactivated in embryonic stem cells by conventional gene targeting as shown in Fig. 1a. Three targeted embryonic stem-cell lines were obtained and one of them produced the germline chimaera that became the founder of the *Id2*-mutant mice (Fig. 1b). No differences were detected between heterozygous mice and their *Id2*^{+/+} littermates at any stage of life. The genotype ratio of born litters from heterozygous intercrosses was 1.0:2.2:0.9 for +/+:-/-:-/- ($n = 462$), indicating mendelian transmission and no embryonic lethality. Northern blot analysis of RNA purified from a newborn liver revealed that the amount of the *Id2* messenger RNA expressed by heterozygous mice is half of that of wild-type mice and the mRNA is not detectable in homozygous mice (Fig. 1c), confirming complete inactivation of the *Id2* gene.

General appearance of *Id2*^{-/-} mice at birth was indistinguishable from that of pups with the other two genotypes, indicating that the Id2 is dispensable in mice during embryogenesis (Fig. 2a, b). However, the *Id2*^{-/-} mice gradually showed retarded growth (Fig. 2a, c) and one quarter of them died in the neonatal period ($n = 79$). Many of the dead *Id2*^{-/-} pups showed distended jejunum and rostral half of ileum, resembling megacolon, and small caecum (data not shown). These anomalies may be related to the retarded growth and neonatal lethality.

In addition, the *Id2*^{-/-} mice lack Peyer's patches—the secondary lymphoid organs that are normally identified as protruded nodes on the antimesenteric side of the intestine from the duodenum to the rectum in mice. Moreover, lymph nodes, including the mesenteric lymph node, are systemically missing (Fig. 3a–d). Lymphatic vessels, in contrast, seem to be normal (Fig. 3c, d, and data not

shown). To determine whether lymphocytes have a normal homing activity to these secondary lymphoid organs, *Id2*^{-/-} bone-marrow cells were transferred into lethally irradiated wild-type mice and the lymphocyte populations (T and B cells) in lymph nodes and Peyer's patches were analysed by flow cytometric analysis. The *Id2*^{-/-} bone-marrow cells successfully populated the secondary lymphoid organs in the same manner as wild-type bone-marrow cells (data not shown), demonstrating normal homing activity of the *Id2*^{-/-} lymphocytes. Consistent with this, adhesion molecules such as 1-selectin and integrins, which have been thought to participate in lymphocyte homing, were detected on splenocytes and peripheral blood lymphocytes of *Id2*^{-/-} mice, similarly to those of control mice (data not shown). In addition, the transfer experiment of wild-type bone-marrow cells into the *Id2*^{-/-} mice failed to induce the generation of Peyer's patches and lymph nodes (data not shown). These observations indicate that the lymphoplasia found in the *Id2*^{-/-} mice is derived from a developmental defect in the generation of secondary lymphoid organs, not from a defect in homing activity of lymphocytes.

For the Peyer's patches, an immunohistochemical analysis has identified three steps during development: appearance of patches consisting of VCAM-1-positive cells at embryonic day 15.5 (E15.5); accumulation of interleukin-7 receptor α (IL-7R α)-positive or

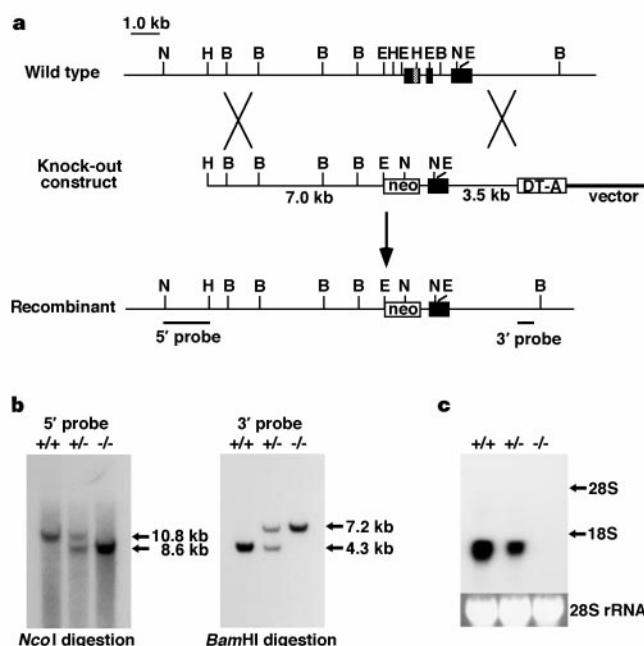


Figure 1 Generation of *Id2*-null mutant mice by targeted inactivation. **a**, Strategy for inactivation of the *Id2* gene. The structure of the wild-type allele, targeting construct and the inactivated allele are illustrated at the top, middle and bottom, respectively. The 2.2-kb genomic sequence in the *Id2* gene locus including the 1.0-kb promoter region was replaced with a neomycin-resistant gene. Filled and hatched boxes indicate exons and the helix-loop-helix region in the coding region of the *Id2*, respectively. Probes of 5' and 3' used in Southern blot analysis are shown in solid bars under the recombinant allele. DT-A and neo are the diphtheria toxin-A fragment and neomycin-resistant genes, respectively. B, *Bam*HI; E, *Eco*RI; H, *Hind*III; N, *Nco*I. **b**, Southern blot analysis of genomic DNA of offspring derived from a heterozygous intercross. *Nco*I-digested genomic DNA was hybridized with the 5' probe and yielded 10.8-kb and 8.6-kb bands representing the wild-type and mutated alleles, respectively. With the 3' probe, 4.3-kb and 7.2-kb bands are indicative of the wild-type and targeted alleles, respectively. **c**, Northern blot analysis of *Id2* expression in wild-type, heterozygous and homozygous neonatal livers. The probe used was the full-length mouse *Id2* cDNA. The intensities of *Id2* transcripts are correlated to the gene dosage. The electrophoretic patterns of 28S and 18S ribosomal RNA are indicated on the right.

[†] Department of Neurology and Psychiatry, School of Medicine, Gifu University, Tsukasa-machi 40, Gifu, Japan (S.S.); Integrative Projects, Mitsubishi Kasei Institute of Life Sciences Minami-Ooya 11, 194-8511 Machida-shi, Tokyo, Japan (S.A.).

CD4-positive cells at E16.5; and collection of CD3- or B220-positive cells at E18.5 (ref. 12). To address at which step the $Id2^{-/-}$ mice have a defect, we performed whole-mount immunohistochemistry and found no patches of VCAM-1⁺ in the neonatal intestine (Fig. 3e, f), indicating $Id2^{-/-}$ mice have a defect in the first step of the formation of Peyer's patches. Loss- and gain-of-function experiments in mice have shown that the family of lymphotoxins (LTs) is indispensable for the formation of secondary lymphoid organs⁶⁻¹¹ and can induce lymphoneogenesis¹³. To determine the localization of LT-expressing cells in the digestive tracts, we conducted a whole-mount *in situ* hybridization using *LT-α* and *-β* antisense RNA probes. The $Id2^{+/-}$ intestine at E18.5 displayed several spots that were positive for the *LT-α* expression on their antimesenteric side (Fig. 3g), demonstrating the clusters of *LT-α*-expressing cells in the prospective Peyer's patches. Furthermore, the analysis using the *LT-β* probe revealed a situation similar to that with the *LT-α* probe (Fig. 3g, i). We next applied this method to determine the distribution of LT-expressing cells in the mutant intestines and found no apparent spot-like staining from E16.5 to the neonatal period, in contrast to the controls (Fig. 3h, j). These *in situ* hybridization studies indicate that LT-producing cells are missing *per se* or that LT-producing cells have some intrinsic defect in *LT* gene expression in the $Id2^{-/-}$ intestine. Alternatively, this method may not be sufficient to detect LT-expressing cells scattered throughout the $Id2^{-/-}$ intestine.

Considering the cells expressing *LT* during development, a cell

population of CD4⁺CD3⁻IL-7Rα⁺ has been identified in the embryonic secondary lymphoid organs¹⁴ and in the gut¹⁵, and considered to be important in the development of secondary lymphoid organs^{11,14,15}. To determine whether these cells are present in the $Id2^{-/-}$ mice, cells were prepared from the embryonic intestine and analysed by flow cytometric analysis. The CD4⁺CD3⁻IL-7Rα⁺-cell population was detected in cells of the $Id2^{+/-}$ intestine but barely detectable in those of the $Id2^{-/-}$ intestine (Fig. 3k). Together with the *in situ* hybridization studies, this indicates that the LT-producing cells *per se* are missing in the $Id2^{-/-}$ intestine and suggests that the defective CD4⁺CD3⁻IL-7Rα⁺-cell population is the cause of alymphoplasia in

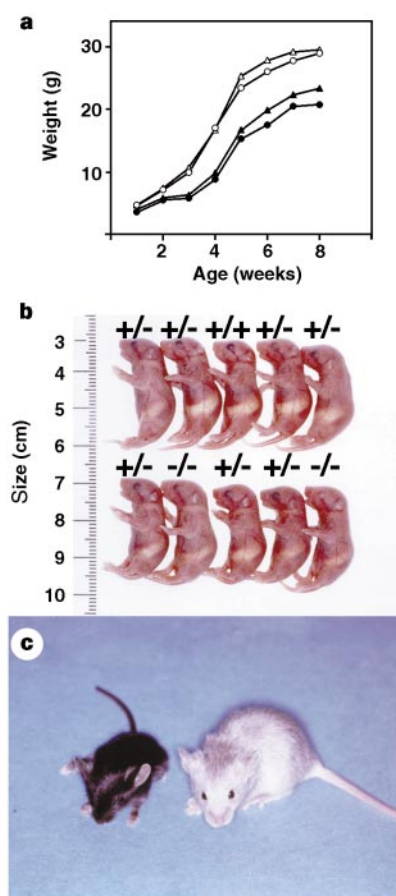


Figure 2 Retarded growth of $Id2^{-/-}$ mice. **a**, Growth curves of $Id2^{+/-}$ and $Id2^{-/-}$ male mice. A representative result is shown. These mice are derived from the same litter and grown in the same cage. Open symbols, $Id2^{+/-}$. Filled symbols, $Id2^{-/-}$. The growth curves of two mice are presented for respective genotypes (circles and triangles). **b**, Appearance at birth of neonates derived from a heterozygous intercross. Genotypes of respective neonates are indicated. **c**, Appearance of $Id2^{-/-}$ mice at one month old. Left and right mice are $Id2^{-/-}$ and $Id2^{+/-}$, respectively, and are littermates kept in the same cage.

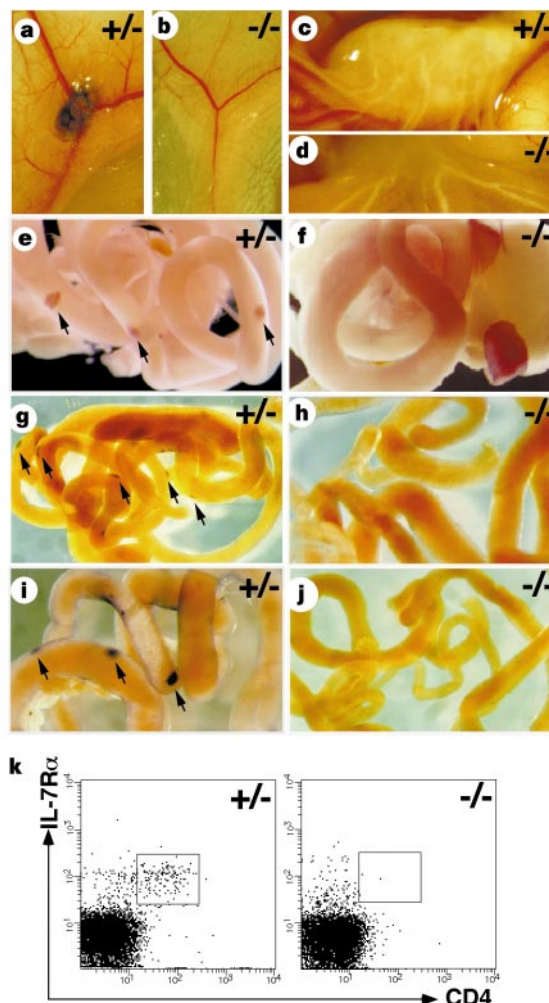


Figure 3 Lymph nodes and Peyer's patches are absent in $Id2^{-/-}$ mice. **a, b**, The right inguinal region one day after injection of black ink into hind footpads. A lymph node collecting black ink is evident in the $Id2^{+/-}$ mouse (**a**) but missing in the $Id2^{-/-}$ mouse (**b**). **c, d**, The mesenteric lymph node is present in the $Id2^{+/-}$ mouse but absent in the $Id2^{-/-}$ mouse. These mice were fed with oil and their mesenteric lymphatic vessels were visualized as white lines. **e, f**, Whole-mount immunohistochemistry of the neonatal intestine on the day of birth for VCAM-1 expression. In the $Id2^{+/-}$ intestine, several patches that are positive for VCAM-1 expression are present (**e**), which correspond to developing Peyer's patches. In contrast, no such VCAM-1-positive region is detected in the $Id2^{-/-}$ intestine (**f**). **g-j**, Whole-mount *in situ* hybridization analysis of the E18.5 intestines. Probes for **g, h** and **i, j** are *LT-α* and *LT-β*, respectively. Several spots positive for expression of *LT-α* and *-β* are visualized in the $Id2^{+/-}$ intestine (**g** and **i**, respectively). The method failed to detect LT-expressing cells in the $Id2^{-/-}$ intestine (**h** and **j**). Arrows in **e, g, i** indicate positive spots. **k**, Flow cytometric analysis of the intestinal cells at E17.5 are presented with the indicated cell-surface markers. The cell population positive for CD4 and IL-7Rα are negative for CD3. The IL-7Rα expression was detected by the monoclonal antibody A7R34 (ref. 30). Cells were prepared from individual embryos and subjected to analysis.

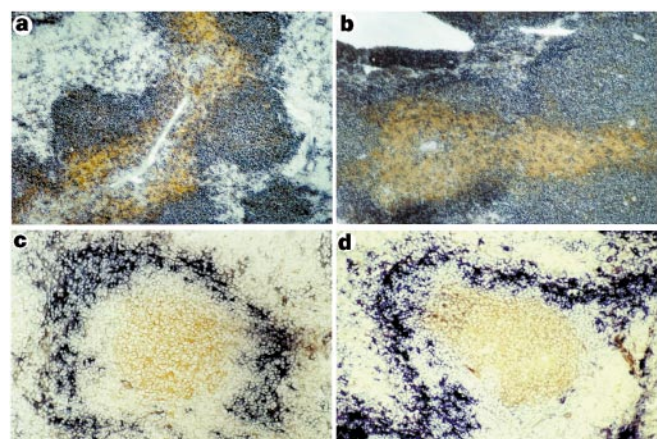
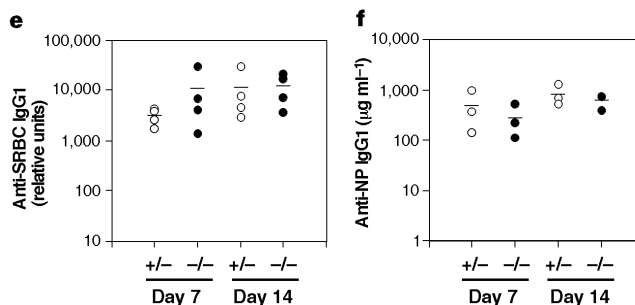


Figure 4 Normal splenic architecture of the *Id2*^{-/-} mice. **a, b**, Distributions of T and B cells in spleen were detected by anti-CD3 and anti-B220 antibodies, and visualized by HRP (brown) and alkaline phosphatase (purple), respectively. Compartmentalization of the two lymphocyte populations is apparent in both *Id2*^{+/-} (**a**) and *Id2*^{-/-} (**b**). **c, d**, Germinal centres and marginal metallophilic macrophages were detected by peanut lectin and a monoclonal antibody, MOMA-1, respectively. Spleens of both *Id2*^{+/-} (**c**) and *Id2*^{-/-} (**d**) show the presence



of germinal centres (brown by HRP) and marginal zones (purple by alkaline phosphatase). **e, f**, T-cell-dependent B-cell response. After immunization of mice, sera were collected at indicated days and specific IgG1 titres were determined by ELISA. Antigens used were SRBC (**e**) and NP-chicken γ -globulin (**a**). Open and filled circles indicate *Id2*^{+/-} and *Id2*^{-/-} titres in respective mice. Each point is the mean of triplicate assay. A horizontal bar shows the mean of each group. $n = 4$ (**e**) and $n = 3$ (**f**).

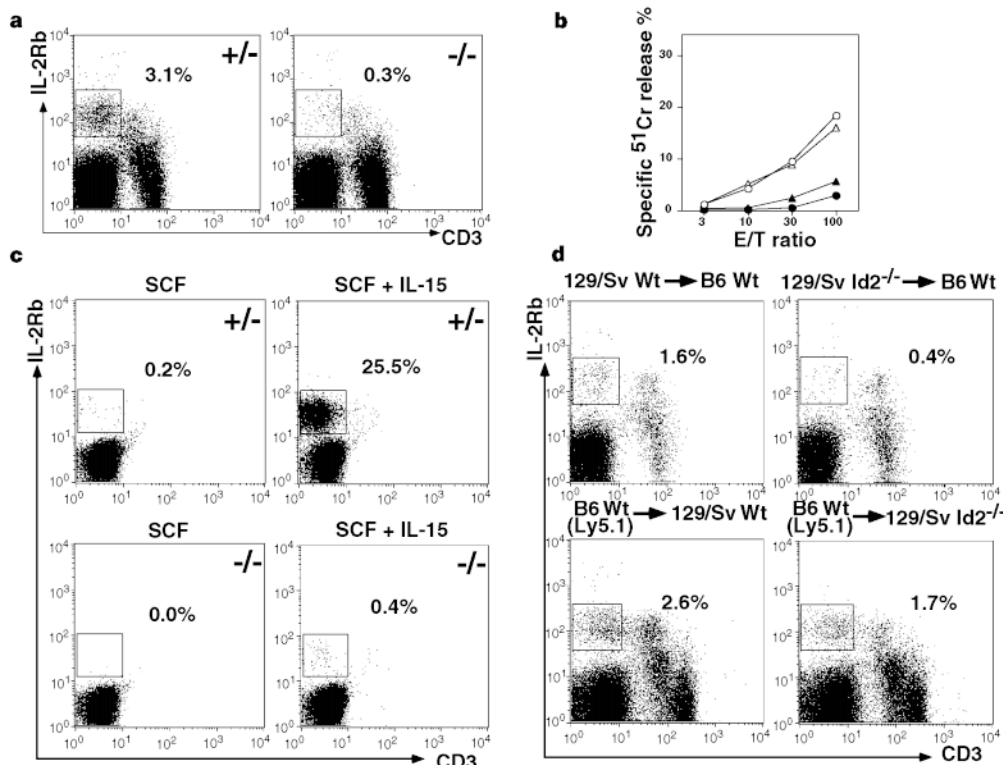


Figure 5 Impaired NK-cell production in the *Id2*^{-/-} mice. **a**, Flowcytometric analysis of the NK-cell population in *Id2*^{+/-} and *Id2*^{-/-} splenocytes. Similar results were obtained with bone-marrow cells and mononuclear cells isolated from the liver. **b**, Cytolytic activity of splenocytes. Splenocytes were subjected to the standard ⁵¹Cr-release assay²⁷ using YAC-1 cells as targets. Open circles and triangles indicate *Id2*^{+/-} and *Id2*^{-/-} splenocytes, respectively. Filled circles and triangles show *Id2*^{-/-} splenocytes prepared from two different homozygous mice. Results are expressed as means of triplicate assays. Similar results were obtained using the RL male 1 cells as target cells and specificity of the lytic activity was confirmed by using cells that were resistant to mouse NK cells. E/T

ratio, effector-to-target ratio. **c**, *In vitro* NK-cell differentiation from bone-marrow cells. Bone-marrow cells cultured with SCF in the presence (right panel) or absence (left panel) of IL-15 were subjected to the flowcytometric analysis. Upper and lower panel show *Id2*^{+/-} and *Id2*^{-/-} cells, respectively. **d**, Flowcytometric analysis of the NK-cell population after bone-marrow transplantation. Host and donor mouse strains are indicated above each panel. WT, wild type. Cells positive for Ly9.1 and Ly5.1 were gated as donor origin and analysed in the upper and lower panels, respectively. The percentage indicated in each panel therefore represents that of NK cells within total donor-derived cells.

the *Id2*^{-/-} mice. Conversely, the *Id2*^{-/-} mice provide evidence that the CD4⁺CD3⁻IL-7Rα⁺-cell population is crucial for the generation of secondary lymphoid organs during development. However, we cannot exclude the possibility that the *Id2*^{-/-} mice have a defect in stromal cells that support accumulation and survival of the CD4⁺CD3⁻IL-7Rα⁺ cells in embryonic intestines.

This generalized lymphoplasia found in the *Id2*^{-/-} mice is reminiscent of a series of targeted null mutants for an LT family and the LT-β receptor⁶⁻¹¹ already mentioned, and *aly*, a spontaneous mutant mouse¹⁶. Besides lymphoplasia, these mice show a disorganized architecture in the spleen, another secondary lymphoid organ, disturbed T-cell and B-cell segregation, and no germinal centre, which is an important place for memory B cells and affinity maturation of immunoglobulins. We then investigated the structure of the *Id2*^{-/-} spleen using immunohistochemical techniques. *Id2*^{-/-} mice showed normal splenic compartments of T cells and B cells, which were identified by anti-CD3 and anti-B220 antibodies, respectively (Fig. 4a, b). Moreover, peanut lectin¹⁷ indicated the presence of distinct germinal centres, and a monoclonal antibody, MOMA-1 (ref. 18), detected marginal metallophilic macrophages that constitute a major component in the marginal zone, features that are similar to those found in the controls (Fig. 4c, d). Distributions of cells expressing VCAM-1 (ref. 19) or MAdCAM-1 (ref. 20) identified no significant difference between the *Id2*^{-/-} and control mice (data not shown).

These results indicate the normal splenic structure of the mutant mice. Furthermore, mice were immunized with sheep red blood cells (SRBC)²¹ or the hapten (4-hydroxy-3-nitrophenyl)acetyl (NP)²², and their specific immunoglobulin-γ-1 (IgG1) levels were analysed to detect T-cell-dependent B-cell responses. The *Id2*^{-/-} mice showed sufficient responses to the antigens, which were similar to those of the control mice, although the mutant mice showed a stronger response to SRBC at day 7 than did the control mice (Fig. 4e, f). In addition, normal serum levels of IgM, IgG and IgA were similar between heterozygous and homozygous mice (data not shown). The *Id2*^{-/-} mice therefore preserve the function of the spleen. This unique situation found in the *Id2*^{-/-} mice indicates that organogenesis is regulated by different mechanisms among secondary lymphoid organs. During an inductive cascade of lymphoid organogenesis^{12,21,23}, the principal LT-producing cell population may be different in the spleen from those in lymph nodes and the

Peyer's patches. Alternatively, compensatory mechanisms may function to complete the inductive events in the spleen.

Despite the lymphoplasia phenotype, the *Id2*^{-/-} mice exhibit a normal peripheral haemogram (data not shown). In addition, the *Id2*^{-/-} thymi are normal in size and show histological structures of cortex and medulla, similar to those of the controls (data not shown). Flow cytometric analysis on surface markers of cells of bone marrow, thymus and spleen also showed no significant difference between the *Id2*^{-/-} and control mice (data not shown), except for the NK cell population. The number of NK cells in the *Id2*^{-/-} splenocytes was approximately 10% of those in the controls (Fig. 5a). The cytolytic activity of the splenocytes was comparable to the proportion of the NK cells found in the *Id2*^{-/-} spleen (Fig. 5b), indicating that the NK cells generated in the *Id2*^{-/-} mice are not functionally affected.

To determine whether the defective NK-cell production is derived from the microenvironment required for NK-cell development²⁴ or from NK-cell precursors, bone marrow cells of *Id2*^{-/-} mice were cultured in the presence of interleukin-15 (IL-15), a key cytokine for NK-cell generation^{24,25}, and analysed. *Id2*^{-/-} bone-marrow cells failed to efficiently produce NK cells compared with wild-type bone-marrow cells (Fig. 5c). Moreover, *Id2*^{-/-} bone-marrow cells that were transferred into wild-type mice poorly differentiated to the NK cells, compared with the efficiency of wild-type bone-marrow cells, whereas wild-type bone-marrow cells that were transferred into the *Id2*^{-/-} mice differentiated to NK cells to a similar extent to those transplanted into wild-type mice (Fig. 5d).

These results show that NK-cell precursors, rather than the microenvironment, have a defect. In accordance with these results, it has been reported that overexpression of Id3 in human CD34⁺ fetal thymocytes promotes NK-cell generation in the fetal thymus organ culture system²⁶, thus providing supporting evidence for the important role of Id proteins in the differentiation process of NK cells, probably from bipotential T/NK progenitor cells^{26,27}.

To investigate whether *Id2* is expressed in the CD4⁺CD3⁻IL-7Rα⁺-cell population and NK cells, these cells were sorted and subjected to polymerase chain reaction with reverse transcription (RT-PCR) analysis. As shown in Fig. 6, both cell types expressed *Id2*, indicating the functional importance of *Id2* in these cells.

What is the relationship between lymphoplasia and defective NK-cell differentiation in the *Id2*^{-/-} mice? One possible explanation might be that the CD4⁺CD3⁻IL-7Rα⁺-cell population is not only a key player in the formation of lymph nodes and Peyer's patches, but also a source of NK cells. This cell population, in fact, can be induced to differentiate to NK cells *in vitro*¹⁴. In addition, the bipotential T/NK progenitors presented by Rodewald partly share cell-surface markers with the CD4⁺CD3⁻IL-7Rα⁺-cell population^{14,27}. The CD4⁺CD3⁻IL-7Rα⁺-cell population, however, is not present in the fetal thymus, whereas the bipotential progenitor cells are detected in the organ. Further studies regarding the origin of the CD4⁺CD3⁻IL-7Rα⁺-cell population would help our understanding of the relation of these cell types. Alternatively, the two phenotypes may not be related to each other in terms of the etiology. *Id2* may play distinct roles in development of the CD4⁺CD3⁻IL-7Rα⁺ cells and in NK-cell differentiation. The current view regarding the function of Id proteins^{2,4} indicates that bHLH factor(s) exist which negatively regulate the generation of the CD4⁺CD3⁻IL-7Rα⁺-cell population and NK cells, probably by promoting differentiation of other cell types from their common precursors. Functional inactivation of such bHLH factor(s) by the *Id2* protein would be essential for development and maintenance of CD4⁺CD3⁻IL-7Rα⁺ and NK cells. Identification of such factor(s) will facilitate elucidation of molecular mechanisms underlying the development of secondary lymphoid organs and NK-cell differentiation. □

Methods

Generation of *Id2*-null-mutant mice. A genomic fragment that covers exons 1

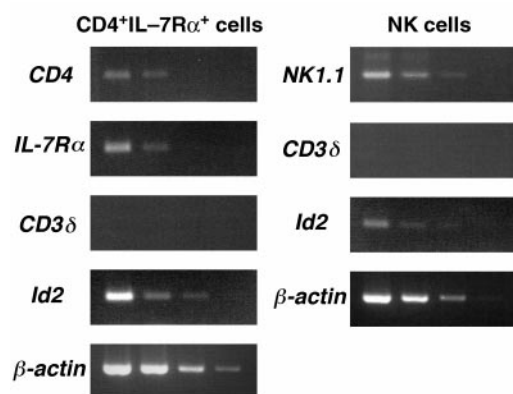


Figure 6 RT-PCR analysis on *Id2* expression in CD4⁺CD3⁻IL-7Rα⁺ and NK cells. The CD4⁺CD3⁻IL-7Rα⁺ and NK cells were isolated from the 15.5-d.p.c. embryonic intestine and adult splenocytes of B6 mice, respectively. The CD4⁺IL-7Rα⁺ population is negative for CD3 expression. *Id2* is expressed in both CD4⁺CD3⁻IL-7Rα⁺ and NK cells. The numbers of cells subjected to RT-PCR are 1,000, 200, 40 and 8 cells from left to right in each panel. Expression of genes indicated on the left was investigated. Positive controls were *CD4* and *IL-7Rα* for CD4⁺CD3⁻IL-7Rα⁺ cells and *NK1.1* for NK cells. The *CD3-δ* chain was a negative control for both cell populations. As an internal control, *β-actin* was included. Without RT, no specific PCR product was detected in each reaction.

and 2 of the mouse *Id2* gene was obtained by PCR, according to the reported sequence³, and used as a probe to isolate a genomic phage clone containing the entire *Id2* gene locus from a 129/Sv mouse genomic library. The targeting construct containing a 7.0-kilobase (kb) *HindIII*–*EcoRI* 5' fragment, the PGKneobpA cassette and a 3.5-kb *BamHI*–*EagI* 3' fragment in pMCIDTpA²⁸ was linearized by digestion with *HindIII* and electroporated into R1 embryonic stem cells. Among 231 G418-resistant clones, 3 were found to have homologous recombination by Southern blot analysis and no additional integration of the targeting construct was identified in these clones with a neo probe. They were used for aggregation with morulae to generate chimaeric mice. One chimaera produced heterozygous offspring by breeding to NMRI or 129/Sv mice. Homozygous mice were obtained by heterozygous intercrosses or by mating of a heterozygous female and a homozygous male.

Histological analysis and immunohistochemistry. For immunohistochemistry tissues were snap-frozen with the TissueTek compound, cut in 8 µm sections and immediately dried. Double staining was done by alkaline phosphatase staining, followed by horseradish peroxidase (HRP). Antibodies and lectin used were anti-CD3 (Caltag), anti-B220 (PharMingen), peanut agglutinin (EY laboratories) and MOMA-1 (Serotec). Anti-CD3 antibody and peanut lectin were biotinylated and the ABC method was used (Vectastain Elite ABC kit, Vector laboratories). Whole-mount immunohistochemistry was done with a monoclonal antibody against VCAM-1 (PharMingen) as described¹².

Flowcytometric analysis. Antibodies used in flowcytometric analysis were against CD3, CD4, CD8, TCRαβ, TCRγδ, B220, CD11b, F4/80, Gr1, c-kit, Ter119 and c-fms. Examined samples were thymocytes, splenocytes, bone-marrow cells and peritoneal exudate cells, with combinations of the antibodies described above. Peripheral blood lymphocytes were analysed for the CD4/CD8 ratio. For NK cells, NK1.1 was not useful to detect the NK-cell population, because of the genetic background of the *Id2*-mutant mice as described above. Instead, expressions of the interleukin-2 receptor β (IL-2Rβ)²⁹ chain and CD16/32 (ref. 27) were used as markers for the NK-cell population in combination with negative CD3 expression. Expression patterns of IL-2Rβ and CD3 are shown in the figures. Preparation of cell suspension from embryonic intestines was done following the method described¹⁵. In brief, embryonic intestines free of mesenteries were treated with dispase (GIBCO-BRL) at 37°C for 20 min, centrifuged, washed with ice-cold phosphate-buffered saline (PBS) containing 5% fetal bovine serum and passed through a 26G needle with a 1-ml syringe several times.

Bone-marrow transplantation. C57BL/6 (B6) mice and the mutant mice with the genetic background of 129/Sv were used for transfer experiments. Bone-marrow cells flushed out from femora were suspended in PBS containing 5% fetal calf serum to produce a concentration of 2.0×10^6 cells per ml. Approximately 1.0×10^6 cells were transferred into lethally irradiated host mice through a tail vein. The dose of irradiation was 9.0 Gy and 7.5 Gy for B6 and 129/Sv, respectively. To detect cells of donor origin, Ly9.1 and Ly5.1 were used in experiments from 129/Sv to B6 mice and from B6 (Ly5.1) to 129/Sv, respectively. Analyses were done 6–7 weeks after transfer.

Whole-mount *in situ* hybridization. Digestive tracts from duodenum to the rectum were isolated and fixed in cold 4% paraformaldehyde/PBS overnight. The samples were subjected to whole-mount *in situ* hybridization with digoxigenin-labelled antisense RNA probes. Polyvinyl alcohol was added into the reaction mixture for alkaline phosphatase to a final concentration of 5%. According to the sequences reported in GenBank, complementary DNAs for mouse *LT-α* and *LT-β* were isolated by RT-PCR and subcloned into appropriate plasmids.

T-cell-dependent B-cell response. Mice of 6–8 weeks old were immunized by intraperitoneal injection of 1×10^8 SRBC or 100 µg of alum-precipitated NP₁₂-chicken γ-globulin. To avoid non-specific stimulation of the immune response, bacterial compound was not included. Blood samples were collected at days 7 and 14 after immunization and sera were separated. SRBC- and NP-specific IgG1 in the samples were determined by the sandwich enzyme-linked immunosorbent assay (ELISA) using plates coated with SRBC or NP₂₂-bovine serum albumin, respectively. Bound antibodies were detected by alkaline phosphatase-conjugated goat anti-mouse IgG1 antibody (Southern Biotechnology Associates). Titres of SRBC-specific IgG1 were calculated by using our standard serum and expressed as relative units to the serum. Quantification of NP-specific IgG1 was done using a standard NP-reactive monoclonal antibody of IgG1 isotype.

Cytolytic activity of NK cells and cell culture. Cytolytic activity of NK cells was determined by measuring ⁵¹Cr release from YAC-1 cells according to the standard method as described²⁷. For splenocytes, mice were injected with polyinosinic-polycytidylic acid 16 h before isolation of the spleen. Bone-marrow cells were cultured in RPMI1640 medium containing 10% fetal calf serum and 100 ng ml⁻¹ stem-cell factor (SCF), with or without 50 ng ml⁻¹ rIL-15 (Genzyme). Flow cytometric and cytolytic analyses were done after culture for 10 days.

RT-PCR analysis. The CD4⁺IL-7Rα⁺ and NK1.1⁺CD3⁻ cells were isolated at 15.5 days *post coitum* (d.p.c.) from the embryonic intestine and adult splenocytes of B6 mice, respectively. The CD4⁺IL-7Rα⁺ population is negative for CD3 expression. Cell sorting was done by a FACS Vantage cell sorter (Becton Dickinson). Oligo-d(T)-primed cDNA prepared from these cells by standard methods was subjected to PCR. Primers used were designed according to the reported sequences in the GenBank database.

Received 28 July 1998; accepted 13 January 1999.

- Weintraub, H. *et al.* The myoD gene family: nodal point during specification of the muscle cell lineage. *Science* **251**, 761–766 (1991).
- Benezra, R. *et al.* The protein Id: a negative regulator of helix-loop-helix DNA binding proteins. *Cell* **61**, 49–59 (1990).
- Sun, X. H., Copeland, N. G., Jenkins, N. A. & Baltimore, D. Id proteins Id1 and Id2 selectively inhibit DNA binding by one class of helix-loop-helix proteins. *Mol. Cell Biol.* **11**, 5603–5611 (1991).
- Norton, J. D., Deed, R. W., Craggs, G. & Sablitzky, F. Id helix-loop-helix proteins in cell growth and differentiation. *Trends Cell Biol.* **8**, 58–65 (1998).
- Iavarone, A. *et al.* The helix-loop-helix protein Id2 enhances cell proliferation and binds to the retinoblastoma protein. *Genes Dev.* **8**, 1270–1284 (1994).
- De Togni, P. *et al.* Abnormal development of peripheral lymphoid organs in mice deficient in lymphotoxin. *Science* **264**, 703–707 (1994).
- Rennert, P. *et al.* Surface lymphotoxin α/β complex is required for the development of peripheral lymphoid organs. *J. Exp. Med.* **184**, 1999–2006 (1996).
- Koni, P. A. *et al.* Distinct roles in lymphoid organogenesis for lymphotoxins α and β revealed in lymphotoxin β-deficient mice. *Immunity* **6**, 491–500 (1997).
- Fütterer, A. *et al.* The lymphotoxin β receptor controls organogenesis and affinity maturation in peripheral lymphoid tissues. *Immunity* **9**, 59–70 (1998).
- Rennert, P. D. *et al.* Lymph node genesis is induced by signaling through the lymphotoxin β receptor. *Immunity* **9**, 71–79 (1998).
- Chaplin, D. & Fu, Y. Cytokine regulation of secondary lymphoid organ development. *Curr. Opin. Immunol.* **10**, 289–297 (1998).
- Adachi, S., Yoshida, H., Kataoka, H. & Nishikawa, S. Three distinctive steps in Peyer's patch formation of murine embryo. *Int. Immunol.* **9**, 507–514 (1997).
- Picarella, D. *et al.* Transgenic tumor necrosis factor (TNF)-α production in pancreatic islets leads to insulinitis, not diabetes. Distinct patterns of inflammation in TNF-α and TNF-β transgenic mice. *J. Immunol.* **150**, 4136–4150 (1993).
- Mebius, R., Rennert, P. & Weissman, I. Developing lymph nodes collect CD4⁺CD3⁺LTB⁺ cells that can differentiate to APC, NK cells, and follicular cells but not T or B cells. *Immunity* **7**, 493–504 (1997).
- Yoshida, H. *et al.* Interleukin-7 receptor α⁺CD3⁺ cells in embryonic intestine induces organizing center of Peyer's patch. *Int. Immunol.* (in press).
- Miyawaki, S. *et al.* A new mutation, *aly*, that induces a generalized lack of lymph nodes accompanied by immunodeficiency in mice. *Eur. J. Immunol.* **24**, 429–434 (1994).
- Rose, M. L. *et al.* Peanut lectin binding properties of germinal centres of mouse lymphoid tissue. *Nature* **284**, 364–366 (1980).
- Kraal, G. & Janes, M. Marginal metallophilic cells of the mouse spleen identified by a monoclonal antibody. *Immunology* **58**, 665–669 (1986).
- Liu, Y. J., Grouard, G., de Bouteiller, O. & Banchereau, J. Follicular dendritic cells and germinal centers. *Int. Rev. Cytol.* **166**, 139–179 (1996).
- Kraal, G. *et al.* Expression of the mucosal vascular addressin, MAdCAM-1, on sinus-lining cells in the spleen. *Am. J. Pathol.* **147**, 763–771 (1995).
- Fu, Y. X. *et al.* Independent signals regulate development of primary and secondary follicle structure in spleen and mesenteric lymph node. *Proc. Natn. Acad. Sci. USA* **94**, 5739–5743 (1997).
- Smith, K. G. C. *et al.* Bcl-2 increases memory B cell recruitment but does not perturb selection in germinal centers. *Immunity* **1**, 803–813 (1994).
- Gonzalez, M. *et al.* The sequential role of lymphotoxin and B cells in the development of splenic follicles. *J. Exp. Med.* **187**, 997–1007 (1998).
- Ogasawara, K. *et al.* Requirement for IRF-1 in the microenvironment supporting development of natural killer cells. *Nature* **391**, 700–703 (1998).
- Carson, W. E. *et al.* Interleukin (IL) 15 is a novel cytokine that activates human natural killer cells via components of the IL-2 receptor. *J. Exp. Med.* **180**, 1395–1403 (1994).
- Heemskerk, M. H. M. *et al.* Inhibition of T cell and promotion of natural killer cell development by the dominant negative helix-loop-helix factor Id3. *J. Exp. Med.* **186**, 1597–1602 (1997).
- Rodewald, H. *et al.* A population of early fetal thymocytes expressing Fc gamma RII/III contains precursors of T lymphocytes and natural killer cells. *Cell* **69**, 139–150 (1992).
- Yagi, T. *et al.* A novel negative selection for homologous recombinants using diphtheria toxin A fragment gene. *Anal. Biochem.* **214**, 77–86 (1993).
- Tanaka, T. *et al.* Selective long-term elimination of natural killer cells *in vivo* by an anti-interleukin 2 receptor β chain monoclonal antibody in mice. *J. Exp. Med.* **178**, 1103–1107 (1993).
- Sudo, T. *et al.* Expression and function of the interleukin 7 receptor in murine lymphocytes. *Proc. Natl Acad. Sci. USA* **90**, 9125–9129 (1993).

Acknowledgements. We thank H. Yoshida, N. Mitato, M. Matsumoto, T. Yamada, S. Takeda, K. Ogasawara and R. Benezra for valuable discussions and suggestions; T. Yagi, T. Neuman, S. Itoharu, H. Yoshida, T. Takemori, M. Miyasaka, T. Nishimura, E. Nakamura, N. Minato, H. Kawamoto and Y. Katsura for providing materials; R. Liebal for supporting the mouse work; J. Krull, Y. Tsuji, S.-H. Park, M. Tanji, T. Kikuchi and M. Hirashima for assistance; and R. Altschäffler for photographic work. This work was supported by the Max-Planck Society, the Ministry of Education, Science and Culture of Japan, and the Yamanouchi Foundation. Y.Y. was supported by the Alexander von Humboldt Foundation, the Jeantet-Foundation and the Human Frontier Science Program.

Correspondence and requests for materials should be addressed to Y.Y. (e-mail: yyokota@virus.kyoto-u.ac.jp).

The head inducer Cerberus is a multifunctional antagonist of Nodal, BMP and Wnt signals

Stefano Piccolo*[‡]§, Eric Agius*[‡], Luc Leyns*[§],
Subha Bhattacharyya*, Horst Grunz*[§],
Tewis Bouwmeester[†] & E. M. De Robertis*

*Howard Hughes Medical Institute and Department of Biological Chemistry,
University of California, Los Angeles, California 90095-1662, USA

[†]EMBL, Meyerhofstrasse 1, 69117 Heidelberg, Germany

[‡]These authors contributed equally to this work

Embryological and genetic evidence indicates that the vertebrate head is induced by a different set of signals from those that organize trunk-tail development^{1–6}. The gene *cerberus* encodes a secreted protein that is expressed in anterior endoderm and has the unique property of inducing ectopic heads in the absence of trunk structures¹. Here we show that the *cerberus* protein functions as a multivalent growth-factor antagonist in the extracellular space: it binds to Nodal, BMP and Wnt proteins via independent sites. The expression of *cerberus* during gastrulation is activated by earlier nodal-related signals in endoderm and by Spemann-organizer factors that repress signalling by BMP and Wnt. In order for the head territory to form, we propose that signals involved in trunk development, such as those involving BMP, Wnt and Nodal proteins, must be inhibited in rostral regions.

The principal activities of injected *cerberus* (*cer*) messenger RNA in *Xenopus* embryos are the inhibition of trunk mesoderm and neuralization of the ectoderm¹. *Xenopus cerberus* protein (Cer) is structurally related to a family of cystine-knot secreted proteins that behave as antagonists of members of the transforming growth factor- β (TGF- β) family^{7,8}. To determine the molecular mechanism of Cer function, we prepared soluble *Xenopus* Cer protein tagged with the Flag epitope. In animal cap cells expressing *cer* mRNA, Cer-Flag was secreted as a soluble protein of relative molecular mass (M_r) 46,000, designated Cer-long (Cer-L); a minor form of M_r 33,000, designated Cer-short (Cer-S), was also produced (Fig. 1a, lane 1). Transfected human 293T cells (Fig. 1a, lane 2) produced almost exclusively the proteolytically processed Cer-S form (Fig. 1b). To test whether Cerberus binds to any of the three main mesoderm-inducing candidate factors, we produced biologically active preparations of activin, Vg1 and Xnr-1 (*Xenopus* nodal-related-1, ref. 9), tagged with haemagglutinin (HA) at the amino terminus (Fig. 1c). In immunoprecipitation experiments Cer-Flag proteins bound Xnr-1 but not activin or Vg1 proteins, displaying a remarkable degree of specificity (Fig. 1d). In the presence of either Cer-L or Cer-S protein, the *Xbra*-inducing activity of Xnr-1 was blocked, but that of activin or Vg1 proteins was not (Fig. 1e). In animal cap assays, we found that at 10^{-9} M, Cer-L and Cer-S proteins titrated Xnr-1 signalling stoichiometrically, indicating that the affinity of the interaction must be in the subnanomolar range (Fig. 1f).

The neutralizing activity of *cer* mRNA can be antagonized by BMP-4 (refs 1,7). We performed co-immunoprecipitation experiments to test whether *Xenopus* Cerberus proteins bind BMPs directly. Cer-L-Flag protein bound BMP-4, with a dissociation constant (K_D) of 0.6 nM (Fig. 1g). This interaction was highly specific, as it could be displaced by a 10-fold excess of the related

molecule BMP-2, but not by TGF- β -1, epidermal growth factor (EGF) or platelet-derived growth factor (PDGF; Fig. 1h, lanes 3–6). In biological assays, Cer-L protein had neutralizing activity (Fig. 1i, lane 3), whereas the proteolytically processed Cer-S protein did not bind BMP-4 (even in 10-fold excess, Fig. 1h, lane 9) and did not induce neural markers (Fig. 1j, lane 4).

An additional property of *cer* mRNA is that it blocks the induction of secondary axes by *Xwnt*-8 mRNA⁴. We found that in animal caps, full-length *cer* mRNA inhibited *Xwnt*-8 but not β -catenin signalling (Fig. 1j), suggesting that Cer-L might antagonize *Xwnt*-8 extracellularly. To test this directly, we incubated *Xwnt*-8-HA protein produced in oocytes with Cer-L-Flag or Cer-S-Flag and immunoprecipitated with anti-Flag antibody. We found that *Xwnt*-8 bound to Cer-L, but not to the Cer-S protein (Fig. 1k, lanes 3–5). Thus, the Cer-L protein is a trifunctional antagonist that can bind Xnr-1, BMP and *Xwnt*-8 in solution.

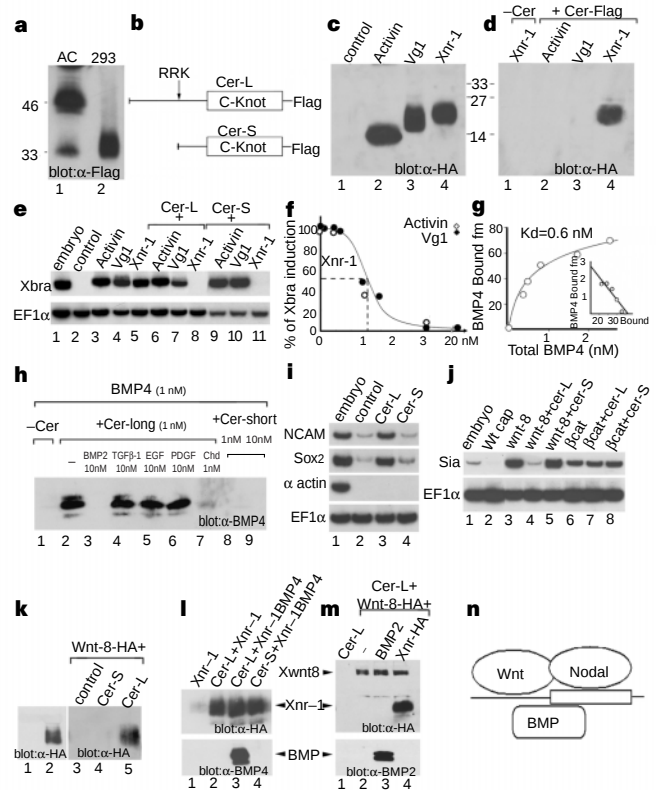


Figure 1 Cerberus protein binds to Xnr-1, BMP-4 and *Xwnt*-8. **a**, Cer-Flag secreted by *Xenopus* animal cap (AC) and cultured 293T cells. **b**, The two Cer protein products. **c**, HA-tagged activin, Vg1 and Xnr-1 secreted by *Xenopus* oocytes. **d**, Xnr-1 is bound specifically by Cer-S-Flag (lanes 2–4) or Cer-L-Flag (not shown). **e**, Cer binding inactivates Xnr-1 signalling. Animal cap explants were treated with oocyte control medium, activin (15 pM), Vg1 (0.3 nM) or Xnr-1 (1 nM) proteins, either alone (lanes 2–5) or together with 5 nM Cer-S (lanes 6–8) or Cer-L (lanes 9–11). **f**, Cer inhibits Xnr-1 with high affinity; 2 nM Xnr-1 was incubated with increasing concentrations of Cer-S (closed circles) or Cer-L (open circles). **g**, Cer-L-Flag binds BMP-4 with a K_D of 0.6 nM. **h**, Binding of Cer-L-Flag to BMP-4 is competed for by BMP-2 but not by TGF- β -1, EGF or PDGF; Cer-S does not bind BMP-4. **i**, Cer-L (10 nM), but not Cer-S (20 nM), is a direct neural inducer of animal cap cells sensitized by brief dissociation-re-aggregation (but not of intact caps). **j**, Cer inhibits *Xwnt*8 but not β -catenin mRNA; induction of *Siamois*, a target of Wnt signalling, was assayed at stage 10+. **k**, Lanes 1, 2, soluble *Xwnt*-8-HA protein is secreted by mRNA-injected oocytes. Lanes 3–5, binding of *Xwnt*-8-HA (5 nM) to Cer-Flag proteins (10 nM). **l**, **m**, The *Xwnt*-8-HA, Xnr-1-HA and BMP binding sites in Cer-L are independent. Cer-L-Flag (2 nM) was bound to Xnr-1 (1 nM) or *Xwnt*-8 (2 nM) and competed with BMP-4 (10 nM), BMP-2 (27 nM) or Xnr-1 (8 nM). **n**, Multiple ligand-binding sites on Cerberus.

§ Present addresses: Istituto di Istologia ed Embriologia, Università di Padova, 35100 Padova, Italy (S.P.); Vrije Universiteit Brussels, 1050 Brussels, Belgium (L.L.); Abteilung Zoophysiology, Universität GH Essen, 45117 Essen, Germany (H.G.).

The binding of Xnr-1 to Cer-L-Flag could not be displaced by excess BMP (Fig. 1l) and Xwnt-8 binding was not displaced by excess BMP or Xnr-1 (Fig. 1m). This suggests that Cer-L binds each ligand at independent binding sites (Fig. 1n). In addition, the biochemical studies show that Cer-L is proteolytically processed into the Cer-S protein containing the cystine-knot (Fig. 1b), which retains only the anti-Xnr-1 activity. This proteolytic cleavage could be important in regulating Cerberus function *in vivo*; a regulatory proteolytic step has been described in the case of chordin (Chd), in which active BMP ligands are released from inactive Chd/BMP complexes¹⁰.

Xenopus Cerberus has three distinct inhibitory activities and, as shown below, each one is required for the induction of ectopic head-like structures. To study the phenotypic effects of specifically inhibiting nodal-related signals in the context of the *Xenopus* embryo, we generated a construct encoding a secreted form of Cer-S, designated *cer-S*. Injection of *cer-S* mRNA into single

blastomeres did not induce ectopic heads but instead gave rise to anterior defects in the endogenous head, including cyclopia (not shown). Embryos in which each blastomere was injected at the four-cell stage with *cer-S* mRNA lacked axial structures and the expression of the trunk mesoderm markers *Xbra* and *Xwnt-8* was blocked (Fig. 2a–d). In addition to maintaining a mesodermal fate, nodal-related signals are required for the formation of anterior endoderm: expression of endogenous *cerberus* was blocked by *cer-S* mRNA (Fig. 2e, lane 2) and induced by *Xnr-1* mRNA (Fig. 2f). Interestingly, endodermal patterning by *Xnr-1* mRNA appears to function early in development, as injection of an *Xnr-1* DNA construct expressed after mid-blastula was unable to induce *cerberus* (Fig. 2f, lanes 6–8).

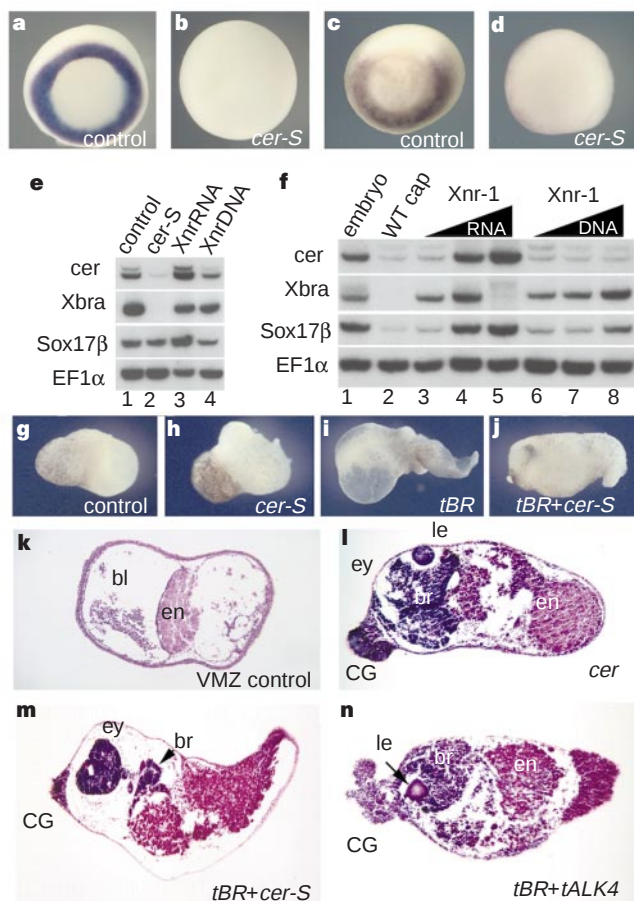


Figure 2 Phenotypic effects of *cer-S* mRNA, an inhibitor of nodal-related signals. Embryos were injected radially at the 4-cell stage. **a–d**, *In situ* hybridization for *Xbra* (**a**, **b**) and *Xwnt-8* (**c**, **d**) in wild type and in embryos injected with 120 pg of *cer-S* mRNA per blastomere. **e**, Anterior endoderm formation requires early nodal-related signals in injected embryos: endogenous *cerberus* transcripts are inhibited by *cer-S* mRNA (200 pg total) and induced by *Xnr-1* mRNA (50 pg). **f**, Dual function for Xnr-1; *Xnr-1* mRNA (6, 12, 24 pg), but not pCS2-*Xnr-1* DNA (20, 40, 80 pg), can induce *cer* expression in animal caps. **g–n**, Head induction by simultaneous repression of BMP and nodal signals in explants of ventral marginal zone (VMZ). **g**, **k**, Wild-type VMZs consisting histologically of endoderm (en) and blood (bl). **h**, *cer-S* mRNA (150 pg); **i**, *tBR* mRNA (600 pg); **j**, **m**, co-injection of *tBR* mRNA (600 pg) and *cer-S* mRNA (150 pg) generates head-like structures (76%, $n = 17$), similar to those produced by *cerberus* mRNA (**l**, 50%, $n = 10$), consisting only of anterior structures with cement gland (CG), brain (br) and eye (ey) containing a lens (le). **n**, Blocking BMP and Nodal signalling by *tBR* (600 pg) and *tALK4* (1.5 ng) RNAs is sufficient to induce head-like structures.

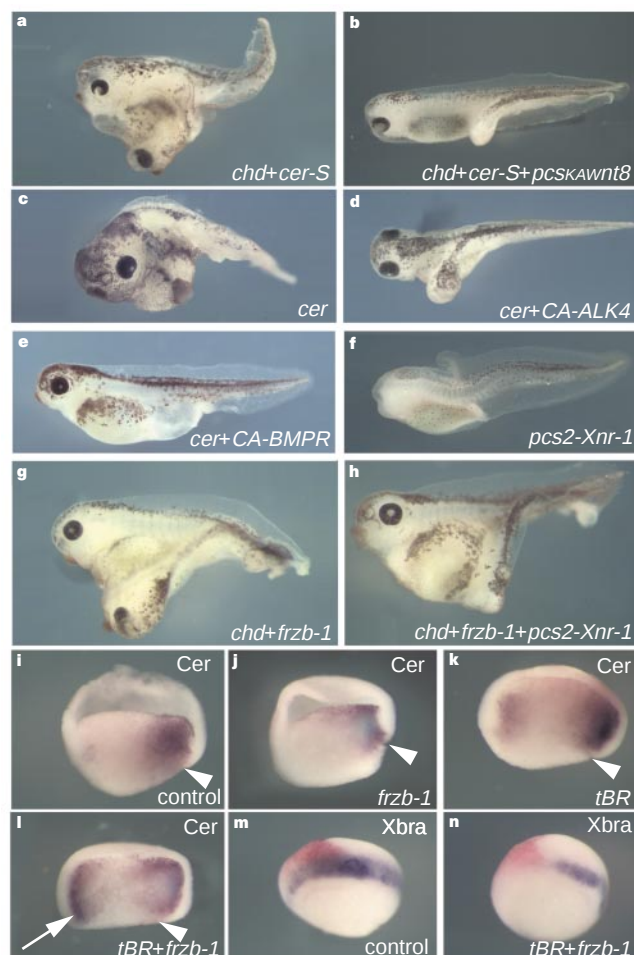


Figure 3 Head induction requires the triple inhibition of nodal-related, Wnt and BMP signals. Embryos were injected into one ventral blastomere with the indicated mRNAs or DNAs at the 4-cell stage. **a**, Ectopic head induction by a combination of *chd* (50 pg) and *cer-S* (200 pg) mRNA (58%, $n = 115$). **b**, pCSKA-*Xwnt-8* DNA (50 pg) antagonizes head induction by *chd/cer-S* (1% ectopic heads, $n = 11$). **c**, Ectopic heads formed by injection of 150 pg *cer* mRNA (57%, $n = 66$) are abolished in **d** and **e** by co-injection of 1 ng CA-*ALK4* mRNA (0%, $n = 36$) or 400 pg of CA-*BMPR* mRNA (0%, $n = 21$). **f**, Dorsal injection of pCS2-*Xnr-1* (32 pg) causes head reduction (53%, $n = 60$). **g**, Co-injection of *chd* (25 pg) and *Frzb-1* (200 pg) mRNAs mediates formation of a complete secondary axis with cyclopic head (49%, $n = 175$). **h**, Co-injection of 16 pg of pCS2-*Xnr-1* DNA inhibits head induction by *chd/Frzb-1* (5%, $n = 133$). **i–n**, Simultaneous repression of BMP and Wnt signalling leads to ectopic activation of *cer* in endoderm and concomitant downregulation of *Xbra* in mesoderm. Embryos were cut sagittally in two halves after fixation and before whole-mount hybridization. Arrowhead indicates the dorsal blastopore lip. *Frzb-1* mRNA (300 pg) synergizes with *tBR* mRNA (800 pg) in upregulating *cer* in the ventral endoderm (arrow). Note in **n** the inhibition of *Xbra* in the injected area, correlating with the activation of *cer* (an anti-Nodal agent) observed in **i**.

Taken together, these gain- and loss-of-function studies support the view that nodal-related signals have a central role in the patterning of anterior endoderm and in trunk mesoderm formation in *Xenopus*, in agreement with genetic studies in other vertebrates^{11–17}.

Since the anti-nodal activity of *cer-S* was not sufficient to induce ectopic head-like structures, we next tested it in combination with anti-BMP reagents in explants of ventral marginal zone (VMZ), which normally develops into postero-ventral mesoderm containing blood (Fig. 2g, k). Microinjection of *cer-S* or of *tBR*, a dominant-negative BMP receptor mRNA⁶, did not cause head formation (Fig. 2h, i). However, injection of combined *cer-S* and *tBR* produced a striking change in VMZ fate into head-like structures (Fig. 2j, m). The explants contained a large cyclopic eye (often with crystalline lens), brain, cement gland and endoderm, but no notochord or somites, indicating the generation of head organizer activity. Indistinguishable head-like structures were formed by injection of full-length *cer* mRNA (Fig. 2l). Co-injection of a dominant-negative putative nodal receptor¹⁸, *tALK4*, with *tBR* also induced head-like structures (Fig. 2n), showing that the *cer* constructs do not have an instructive role. Although these experiments in VMZs show that simultaneous inhibition of nodal-related and BMP signalling is sufficient for head specification, they do not indicate whether the anti-Wnt activity of *cerberus* is required for head formation. Indeed, in VMZs the main Wnt, *Xwnt-8*, is inhibited at the gene-expression level in embryos injected with *cer-S* (Fig. 2d) or *tBR* mRNA¹⁹.

To test whether triple inhibition of Wnt, nodal-related and BMP signalling is required for ectopic head formation, we carried out epistatic combinatorial experiments in the context of the whole embryo. Co-injection of *chd* (anti-BMP) and *cer-S* (anti-nodal) into a ventral blastomere resulted in ectopic heads containing a cyclopic eye (58%, $n = 115$); this induction was blocked by co-injection of *Xwnt-8* DNA (Fig. 3a, b). Similarly, selective inactivation of the anti-nodal activity of full-length *cer* mRNA by co-injection of a constitutively active ALK-4 receptor¹⁸, even at doses that only weakly promote trunk mesoderm formation, abolished ectopic head induction by *cer* (Fig. 3c, d). Mimicking BMP signalling by co-injecting a constitutively active BMP receptor construct also antagonized ectopic head induction (Fig. 3e). These epistatic experiments indicate that all three inhibitory activities of *cerberus* are required for the formation of ectopic head structures.

Niehrs and co-workers have shown that simultaneous inhibition

of BMP and Wnt signalling is sufficient to induce secondary axes containing heads with cyclopic eyes⁴. This is puzzling, for this head induction would occur without apparent inhibition of nodal-related signals. We confirmed these results using co-injection of *chd*¹⁰ and *Frzb-1* (ref. 20) mRNAs, which encode secreted inhibitors specific for BMP and Wnt proteins, respectively (Fig. 3g). Head formation by *chd* and *Frzb-1* was suppressed by co-injection of *Xnr-1* DNA (Fig. 3h and 3f), suggesting that inhibition of nodal signalling could also be involved in this experimental situation, perhaps by activating transcription of *cerberus*. As shown in Fig. 3i–l, co-injection of *Frzb-1* and *tBR* mRNA led to strong ectopic expression of *cerberus* in ventral endoderm at levels comparable to those seen with endogenous *cerberus* mRNA (Fig. 3l, arrow), with concomitant inhibition of the trunk mesodermal marker *Xbra* near the site of injection (Fig. 3m, n).

In mice, the anterior visceral endoderm (AVE) has a fundamental role in head formation^{2,14,16,21} and is the topological equivalent of the anterior endoderm that expresses *cerberus* in *Xenopus*^{2,22}. Robertson and colleagues have shown that *nodal* activity in the AVE is required for head formation in chimaeric embryos¹⁴, whereas we report here that inhibition of nodal-related signalling by *cerberus* is required for head formation. These paradoxical observations can be reconciled by the proposal that, in mice, *nodal* signalling^{14–17} may be required in the AVE to express secreted factors that antagonize the formation of mesoderm in the overlying epiblast^{16,17}, in which the head forms. In *Xenopus*, we have found evidence that *cerberus* expression requires an early nodal-related function mimicked by injection of *Xnr-1* mRNA, but not DNA (Fig. 2e, f). This early signal, which correlates with the expression of *Xnr-1* in endoderm at late blastula⁹, leads to the expression of a multivalent antagonist, Cerberus, which inhibits the function of nodal-related signals in rostral regions at later stages, preventing trunk mesoderm from forming in the head field. As shown in Fig. 4, the expression of *cerberus* in anterior endoderm at the mid-gastrula stage¹ results from earlier signals provided by TGF- β proteins expressed in the endoderm^{9,23} as well as by organizer-specific secreted inhibitors of the Wnt^{5,20,21} and BMP⁶ pathways. In this view, the secretion of Cerberus into the extracellular space would be an important downstream event that would lock the head-organizer programme in place by simultaneously antagonizing three signalling pathways involved in trunk formation, thereby restricting the trunk territory to the posterior part of the body. □

Methods

DNA constructs. To generate HA-tagged soluble *Xnr-1*, *Vg1*, activin and *Xwnt-8*, we designed a secreted protein-expression cassette vector that contains the pre-pro region and proteolytic cleavage site of activin- β B²⁴ followed by the sequence RGL(YPYDVPDYA)LE (pCS2–*proAct*–HA). The two amino acids (LE) after the HA tag were inserted to provide an *XhoI* cloning site. The mature regions of signalling proteins were generated by the PCR reaction using primers flanked by *XhoI* and *XbaI* sequences and subcloned into the HA expression cassette. Mature signalling sequences started at residue C259 of activin- β B, S247 of *Vg1*, N286 of *Xnr-1* and V26 of *Xwnt-8*. pCS2–*cer-S* was generated by PCR, deleting residues R26 to A116 of *cer*.

Protein binding. Soluble *Xwnt-8*, *Xnr-1*, *Vg1* and activin were secreted by manually defolliculated *Xenopus* oocytes injected with 50 ng of mRNA and incubated in 5–10 μ l per oocyte of OR2 medium for 2 days²⁴. *Cer-S* was prepared by transfection of 293T cells¹⁰; *Cer-L* was prepared from animal caps of embryos injected with 1 ng *cer-Flag* mRNA. Typically, 10 caps were dissociated at stage 9 and incubated in 100 μ l of Ca–Mg-free medium for 3 h. For immunoprecipitations, protein mixtures were incubated on ice (2 h) in 20 mM Tris–HCl pH 7.5, 150 mM NaCl, 1 mM CaCl₂, 1 mM MgCl₂, 2% glycerol, 0.1% BSA and 0.1% each of the detergents Triton X-100, CHAPS and octylglucoside (Pierce). Protein-A beads pre-bound with anti-Flag polyclonal antibody (Santa Cruz Inc.) were then added and samples incubated with end-over-end rotation for 1 h at 4 °C. After three 5-min washings in the same buffer, protein complexes were analysed by electrophoresis.

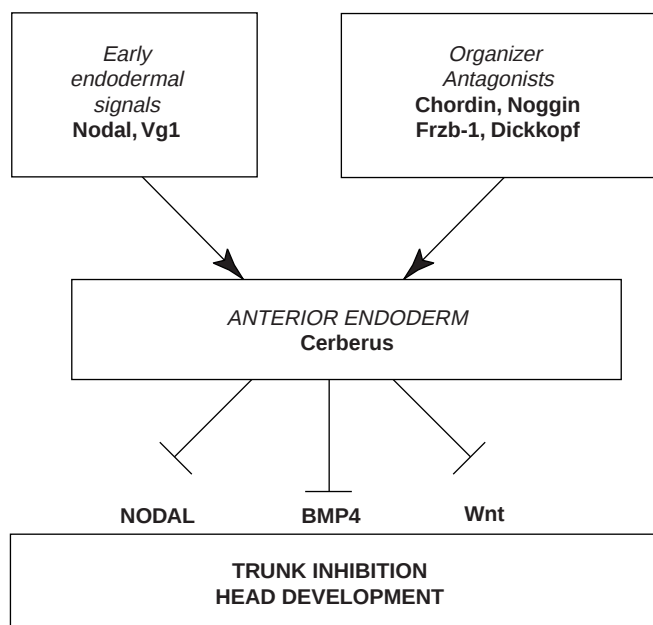


Figure 4 Model of the formation and function of anterior endoderm in *Xenopus* head development.

Received 23 November 1998; accepted 11 January 1999.

1. Bouwmeester, T., Kim, S. H., Sasai, Y., Lu, B. & De Robertis, E. M. Cerberus is a head-inducing secreted factor expressed in the anterior endoderm of Spemann's organizer. *Nature* **382**, 595–601 (1996).
2. Thomas, P. & Beddington, R. Anterior primitive endoderm may be responsible for patterning the anterior neural plate in the mouse embryo. *Curr. Biol.* **6**, 1487–1496 (1996).
3. Shawlot, W. & Behringer, R. R. Requirement for *Lim1* in head-organizer function. *Nature* **374**, 425–430 (1995).
4. Glinka, A., Wu, W., Onichtchouk, D., Blumenstock, C. & Niehrs, C. Head induction by simultaneous repression of Bmp and Wnt signalling in *Xenopus*. *Nature* **389**, 517–519 (1997).
5. Glinka, A. *et al.* Dickkopf-1 is a member of a new family of secreted proteins and functions in head induction. *Nature* **391**, 357–362 (1998).
6. Harland, R. & Gerhart, J. Formation and function of Spemann's organizer. *Annu. Rev. Cell Dev. Biol.* **13**, 611–667 (1997).
7. Hsu, D. R., Economides, A. N., Wang, X., Eimon, P. M. & Harland, R. M. The *Xenopus* dorsalizing factor gremlin identifies a novel family of secreted proteins that antagonize BMP activities. *Mol. Cell* **1**, 673–683 (1998).
8. Belo, J. A. *et al.* Cerberus-like is a secreted factor with neuralizing activity expressed in the anterior primitive endoderm of the mouse gastrula. *Mech. Dev.* **68**, 45–57 (1997).
9. Jones, C. M., Kuehn, M. R., Hogan, B. L. M., Smith, J. C. & Wright, C. V. E. Nodal-related signals induce axial mesoderm and dorsalize mesoderm during gastrulation. *Development* **121**, 3651–3662 (1995).
10. Piccolo, S. *et al.* Cleavage of Chordin by Xoloid metalloprotease suggests a role for proteolytic processing in the regulation of Spemann organizer activity. *Cell* **91**, 407–416 (1997).
11. Feldman, B. *et al.* Zebrafish organizer development and germ-layer formation require nodal-related signals. *Nature* **395**, 181–185 (1998).
12. Sampath, K. *et al.* Induction of the zebrafish ventral brain and floor plate requires cyclops/nodal signalling. *Nature* **395**, 185–189 (1998).
13. Rebagliati, M. R., Toyama, R., Haffter, P. & Dawid, I. B. *cyclops* encodes a nodal-related factor involved in midline signaling. *Proc. Natl Acad. Sci. USA* **95**, 9932–9937 (1998).
14. Varlet, I., Collignon, J. & Robertson, E. J. Nodal expression in the primitive endoderm is required for the specification of the anterior axis during mouse gastrulation. *Development* **124**, 1033–1044 (1997).
15. Conlon, F. L. *et al.* A primary requirement for *nodal* in the formation and maintenance of the primitive streak in the mouse. *Development* **120**, 1919–1928 (1994).
16. Waldrup, W. R., Bikoff, E. K., Hoodless, P. A., Wrana, J. L. & Robertson, E. J. Smad2 signaling in extraembryonic tissues determines anterior-posterior polarity of the early mouse embryo. *Cell* **92**, 797–808 (1998).
17. Nomura, M. & Li, E. Smad2 role in mesoderm formation, left-right patterning and craniofacial development. *Nature* **393**, 786–790 (1998).
18. Chang, C., Wilson, P. A., Mathews, L. S. & Hemmati-Brivanlou, A. A *Xenopus* type I activin receptor mediates mesodermal but not neural specification during embryogenesis. *Development* **124**, 827–837 (1997).
19. Hoppler, S. & Moon, R. T. *BMP-2/4* and *Wnt-8* cooperatively pattern the *Xenopus* mesoderm. *Mech. Dev.* **71**, 119–129 (1998).
20. Leyns, L., Bouwmeester, T., Kim, S. H., Piccolo, S. & De Robertis, E. M. Frzb-1 is a secreted antagonist of Wnt signaling expressed in the Spemann Organizer. *Cell* **88**, 747–756 (1997).
21. Rhinn, M. *et al.* Sequential roles for *Otx2* in visceral endoderm and neuroectoderm for forebrain and midbrain induction and specification. *Development* **125**, 845–856 (1998).
22. Bouwmeester, T. & Leyns, L. Vertebrate head induction by anterior primitive endoderm. *BioEssays* **19**, 855–863 (1997).
23. Henry, G. L. & Melton, D. A. *Mixer*, a homeobox gene required for endoderm development. *Science* **281**, 91–96 (1998).
24. Kessler, D. S. & Melton, D. A. Induction of dorsal mesoderm by soluble, mature Vg1 protein. *Development* **121**, 2155–2164 (1995).

Acknowledgements. We thank L. S. Mathews, N. Ueno, R. Moon, D. Melton and K. Cho for gifts of plasmids, K. Masuhara for anti-BMP4 mAb, Genetics Institute for BMPs, and members of the De Robertis laboratory for comments. S.P., E.A., S.B. and H.G. were supported by Telethon, ARC, MSTP and DFG, respectively. This work was supported by a grant from the NIH. E.M.D.R. is a Howard Hughes Medical Institute investigator.

Correspondence and requests for materials should be addressed to E.M.D.R.

Inhibition of transforming growth factor- β /SMAD signalling by the interferon- γ /STAT pathway

Luis Ulloa, Jacqueline Doody & Joan Massagué

Cell Biology Program and Howard Hughes Medical Institute, Memorial Sloan-Kettering Cancer Center, New York, New York 10021, USA

Transforming growth factor- β (TGF- β) and interferon- γ (IFN- γ) have opposite effects on diverse cellular functions^{1–5}, but the basis for this antagonism is not known⁶. TGF- β signals through a receptor serine kinase that phosphorylates and activates the transcription factors Smads 2 and 3 (refs 7, 8), whereas the IFN- γ receptor and its associated protein tyrosine kinase Jak1 mediate phosphorylation and activation of the transcription factor Stat1

(refs 6, 9, 10). Here we present a basis for the integration of TGF- β and IFN- γ signals. IFN- γ inhibits the TGF- β -induced phosphorylation of Smad3 and its attendant events, namely, the association of Smad3 with Smad4, the accumulation of Smad3 in the nucleus, and the activation of TGF- β -responsive genes. Acting through Jak1 and Stat1, IFN- γ induces the expression of Smad7, an antagonistic SMAD^{11,12}, which prevents the interaction of Smad3 with the TGF- β receptor. The results indicate a mechanism of transmodulation between the STAT and SMAD signal-transduction pathways.

The interplay of different signalling pathways is key to providing an integrated response to total signal inputs to the cell. To investigate the possible effects of IFN- γ /STAT signalling on the TGF- β signal-transduction pathway, we used a genetically defined cell system that has been used previously to delineate IFN- γ signalling

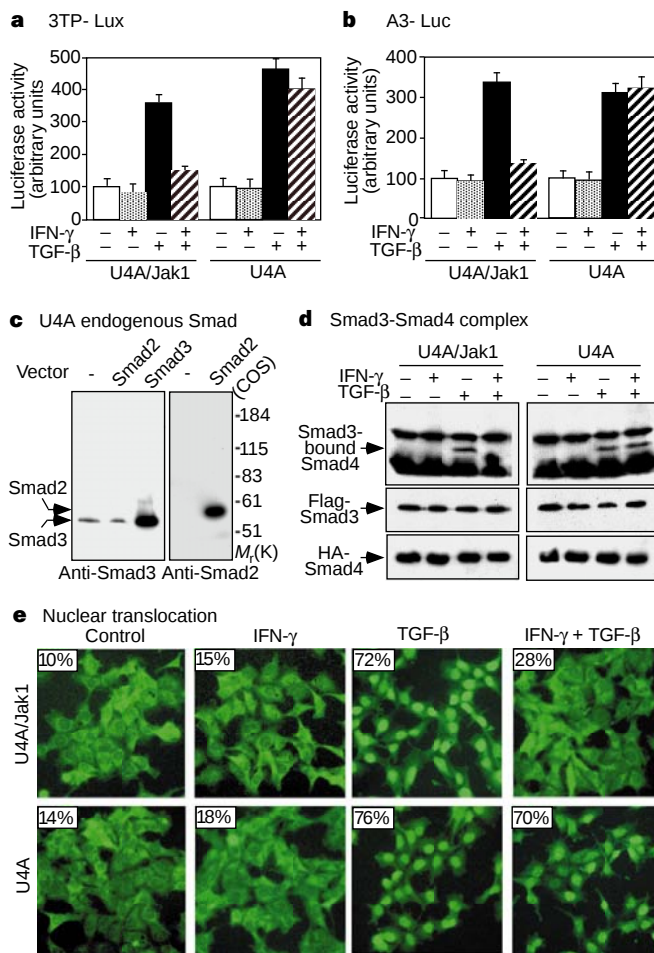


Figure 1 Inhibition of SMAD activation and transcriptional responses by IFN- γ . **a**, **b**, Responses of reporter genes (3TP-Luc and A3-Luc) to TGF- β in U4A and U4A/Jak1 cells, and inhibition of these responses by IFN- γ . Values are averages of triplicate determinations $\pm$ s.d. **c**, U4A/Jak1 cells transfected with SMAD vectors or empty vector were subjected to immunoblotting with Smad3-specific antibodies (left panel). Lysates from U4A/Jak1 cells (right panel, left lane) and Smad2-transfected COS cells (right panel, right lane) were subjected to anti-Smad2 immunoblotting. **d**, U4A/Jak1 and U4A cells transiently co-transfected with Flag-tagged Smad3 and haemagglutinin (HA)-tagged Smad4 were incubated with IFN- γ for 30 min and then TGF- β for 1 h, as indicated. Anti-Flag immunoprecipitates were subjected to anti-HA immunoblotting. Flag-Smad3 and HA-Smad4 expression levels were monitored by anti-epitope immunoblotting. **e**, Smad3 immunofluorescence in cells incubated with IFN- γ and/or TGF- β . The percentages of cells showing predominantly nuclear staining are indicated.

through Jak1 and Stat1 (ref. 6). The U4A cell line is a Jak1-deficient mutant whose responsiveness to IFN- γ can be restored by stable expression of exogenous Jak1, as in the stably transfected subclone U4A/Jak1 (ref. 13). In both cell lines, TGF- β activated 3TP-Lux, a reporter construct containing TGF β -responsive elements from the plasminogen-activator inhibitor-1 and collagenase I genes that control the expression of luciferase¹⁴ (Fig. 1a). IFN- γ inhibited this response to TGF- β in U4A/Jak1 cells but not in Jak1-deficient U4A (Fig. 1a). Similar results were obtained with A3-Luc, a reporter construct whose activation by SMADs is well characterized. A3-Luc contains elements from the *Xenopus Mix.2* gene that are activated by direct binding of a TGF β - or activin-induced transcriptional complex containing Smad2 (or Smad3), Smad4 and the DNA-binding factor Fast1 (refs 15, 16 and F. Liu and J.M., unpublished observations). Transfection of Fast1 endowed U4A and U4A/Jak1 cells with the ability to activate A3-Luc in response to TGF- β (Fig. 1b). IFN- γ inhibited this response in U4A/Jak1 cells but not in U4A cells (Fig. 1b).

U4A/Jak1 cells express Smad3 and no detectable Smad2 (Fig. 1c). Because the levels of Smad3 in U4A/Jak1 cells were not changed by treatment with IFN- γ (see Fig. 5a), we reasoned that the inhibitory effects of IFN- γ on SMAD-dependent gene responses must occur at the level of Smad3 activation. Key events in this process include the association of receptor-activated Smad3 with Smad4 and the translocation of this complex to the nucleus^{7,8}. Formation of a Smad3-Smad4 complex (Fig. 1d) and nuclear accumulation of endogenous Smad3 in response to TGF- β (Fig. 1e) were inhibited by IFN- γ in U4A/Jak1 cells but not in U4A cells. IFN- γ alone did not have any discernible effect on the interactions or subcellular distribution of Smad3. IFN- γ was more effective against 20 pM

TGF- β (Fig. 1e) than against 100 pM TGF- β (data not shown), indicating that the extent of Smad3 activation may be determined by the balance of TGF- β and IFN- γ signals. These results argue that IFN- γ , signalling through Jak1, inhibits the TGF β -induced activation and nuclear accumulation of Smad3.

Accumulation of SMADs in the nucleus is regulated by positive and negative signals. Accumulation is induced by receptor-mediated phosphorylation of SMADs at their carboxy terminus^{7,8} and can be inhibited by MAP-kinase-mediated phosphorylation of SMADs in their central region¹⁷ (M. Kretzschmar, J. D. and J. M., unpublished observations). IFN- γ increases activity of the MAP kinase Erk in U4A/Jak1 cells¹⁸. IFN- γ did cause a small increase in the phosphorylation level of exogenous Flag-Smad3 (Fig. 2a) or endogenous Smad3 (see below) in [³²P]orthophosphate-labelled U4A/Jak1 cells. This increase was rapid and transient, being maximal (on average two- to threefold over basal) 15 min after IFN- γ addition (Fig. 2a). This phosphorylation occurred mainly on serine residues (Fig. 2b) and was inhibited by cell incubation with PD98059 (Fig. 2c), which is an inhibitor of the Erk-activating kinase MEK1 (ref. 19). However, PD98059 did not prevent the IFN- γ -mediated inhibition of the A3-Luc response (Fig. 2d). These results indicate that IFN- γ may induce Smad3 phosphorylation by Erk but that this phosphorylation is not required for IFN- γ -mediated inhibition of TGF- β signalling.

In light of these results, we proposed that IFN- γ specifically inhibits an early step in the TGF β -induced activation of Smad3. One possible mediator of such an effect is Smad7, a member of a SMAD subfamily, known as the 'anti-SMADs', that interfere with TGF- β signalling^{7,8}. Smad7 binds to the TGF β -receptor complex, preventing its interaction with, and phosphorylation of, SMADs^{11,12}.

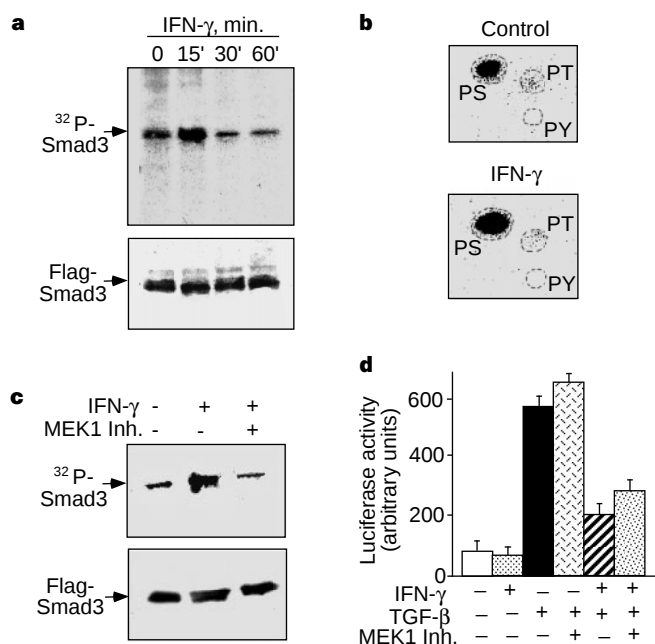


Figure 2 Erk phosphorylation is present but not required for Smad inhibition by IFN- γ in U4A/Jak1 cells. **a**, IFN- γ induces phosphorylation of transfected Flag-Smad3 in U4A/Jak1 cells. Flag-Smad3 levels were monitored by anti-Flag immunoblotting (bottom). **b**, ³²P-labelled Flag-Smad3 from cells treated or not treated with IFN- γ for 15 min was subjected to phosphoamino-acid analysis. Positions shown correspond to phosphoserine (PS), phosphothreonine (PT) and phosphotyrosine (PY). **c**, Smad3 phosphorylation in U4A/Jak1 cells incubated with MEK1 inhibitor for 1 h and IFN- γ for 15 min. **d**, A3-Luc activation in cells incubated with the indicated agents. Values are averages of triplicate determinations $\pm$ s.d.

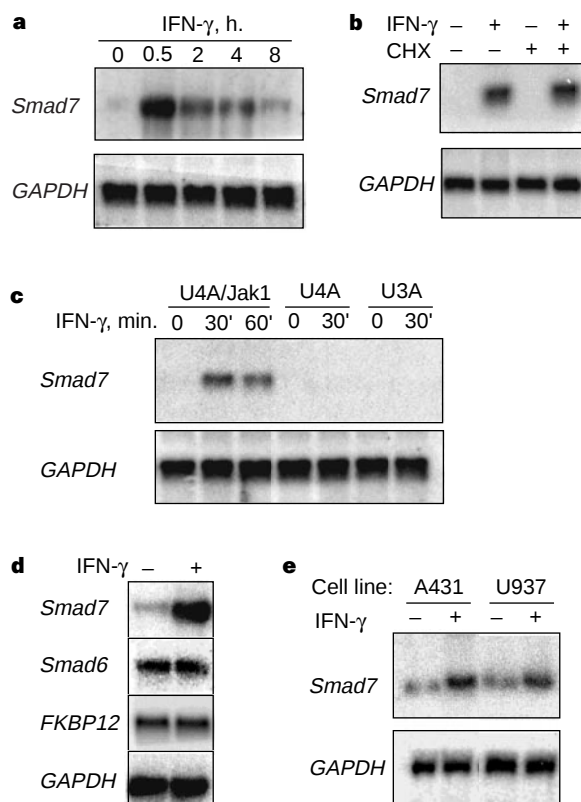


Figure 3 Induction of an antagonistic SMAD by IFN- γ . U4A/Jak1 cells (**a**, **d**) or the indicated cells lines (**c**, **e**) were incubated with IFN- γ for 1 h or the indicated times, and *Smad7*, *Smad6* and *FKBP12* mRNA levels were determined by northern blotting of poly(A)⁺ RNA. To control for mRNA loading, the same blots were probed for *glyceraldehyde-3-phosphate dehydrogenase* (*GAPDH*). In **b**, cells were preincubated with cycloheximide for 1 h before IFN- γ addition.

Indeed, IFN- γ induced the expression of *Smad7* in U4A/Jak1 cells, increasing the messenger RNA levels at least tenfold over the basal level, which was barely detectable (Fig. 3a). This increase was maximal 30 min after IFN- γ addition and still present after 8 h. It was not prevented by the protein-synthesis inhibitor cycloheximide (Fig. 3b), and occurred neither in U4A cells lacking Jak1 nor in U3A cells (Fig. 3c), which lack Stat1 (ref. 20). IFN- γ did not affect the expression of *Smad6*, which is another anti-SMAD^{7,8}, or that of FKBP12, which inhibits TGF β -receptor activation^{7,8} (Fig. 3d). IFN- γ also increased *Smad7* expression in cell types other than U4A cells, which are derived from a human fibrosarcoma¹³. IFN- γ raised *Smad7* expression in human monocytic leukaemia U937 cells, which have opposite responses to IFN- γ and TGF- β in terms of major histocompatibility complex antigen expression and cell adhesion^{3,21}, and in epidermoid carcinoma A431 cells, which have opposite proliferative responses to IFN- γ and TGF- β ^{22,23} (Fig. 3e). These results indicate that *Smad7* induction by IFN- γ is a specific and direct gene response that is mediated by the Jak1/Stat1 pathway in different cell types.

A hallmark of *Smad7* action is its ability to inhibit the association of TGF- β receptors with substrate SMADs^{11,12,24}. To seek evidence of this effect in cells treated with IFN- γ , we labelled the TGF- β type-I receptor (T β RI) and type-II receptor (T β RII) components by crosslinking them to bound ¹²⁵I-labelled TGF- β and measured the association of the labelled receptor complex with endogenous *Smad7* and exogenous Flag-*Smad3*. IFN- γ increased *Smad7* protein levels in U4A/Jak1 cells, as determined by western immunoblotting with anti-*Smad7* antibodies (Fig. 4a). Correlating with this increase, IFN- γ increased the level of *Smad7*-bound receptor complex (Fig. 4a) and decreased the amounts of *Smad3*-bound receptor complex (Fig. 4b). These effects were not observed in U4A cells lacking Jak1. Thus, the type of interference that IFN- γ causes in TGF- β signalling is that which would be expected from the presence of raised *Smad7* levels in the cell^{11,12,24}.

Several observations indicate that the IFN- γ -induced increase in *Smad7* amounts effectively inhibits the ability of the TGF- β receptor to phosphorylate, and thus activate, *Smad3*. A short preincubation with IFN- γ inhibited TGF β -induced phosphorylation of endogenous *Smad3* (Fig. 5a) or transfected *Smad3* (Fig. 5b) in U4A/Jak1 cells down to the level observed with IFN- γ alone (that is, in the absence of TGF- β). This effect of IFN- γ requires Jak1: it was not observed in U4A cells (Fig. 5b). It also requires Stat1, because it was not observed in U3A cells unless they were transfected with a Stat1 vector²⁰ (Fig. 5c). Inhibition of TGF β -induced *Smad3* phos-

phorylation by IFN- γ requires *de novo* protein synthesis, as it was inhibited by cycloheximide (Fig. 5d), and is independent of Erk, as it persisted in the presence of a MEK1 inhibitor (Fig. 5e). Finally, cell treatment with a *Smad7* antisense oligonucleotide that prevented the IFN- γ -mediated increase in *Smad7* amounts (Fig. 5f) also prevented the effect of IFN- γ on *Smad3* phosphorylation (Fig. 5g). The corresponding sense oligonucleotide and a different antisense

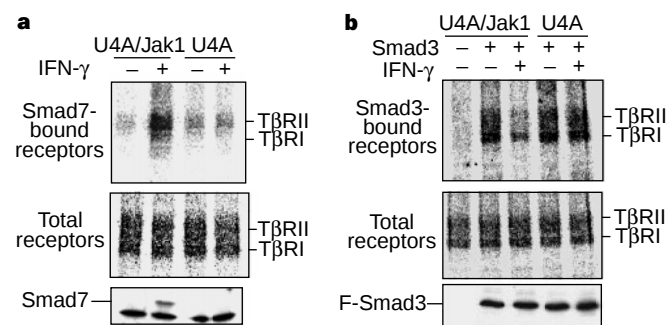


Figure 4 Inhibition of interactions between TGF- β receptors and *Smad3* by IFN- γ . **a**, Receptor interaction with endogenous *Smad7*, as determined by immunoprecipitation with anti-*Smad7* antibodies (top). IFN- γ was added for 30 min. Cell-surface receptors were labelled by crosslinking to bound ¹²⁵I-labelled TGF- β . **b**, Receptor interaction with Flag-*Smad3*, as determined by anti-Flag immunoprecipitation (top). Cells were transfected with Flag-*Smad3* and incubated with IFN- γ for 30 min. **a**, **b**, Aliquots of cell lysates were subjected to SDS-PAGE and autoradiography. (middle panels). Other aliquots were subjected to anti-*Smad7* (**a**) or anti-Flag (**b**) western immunoblotting (bottom panels).

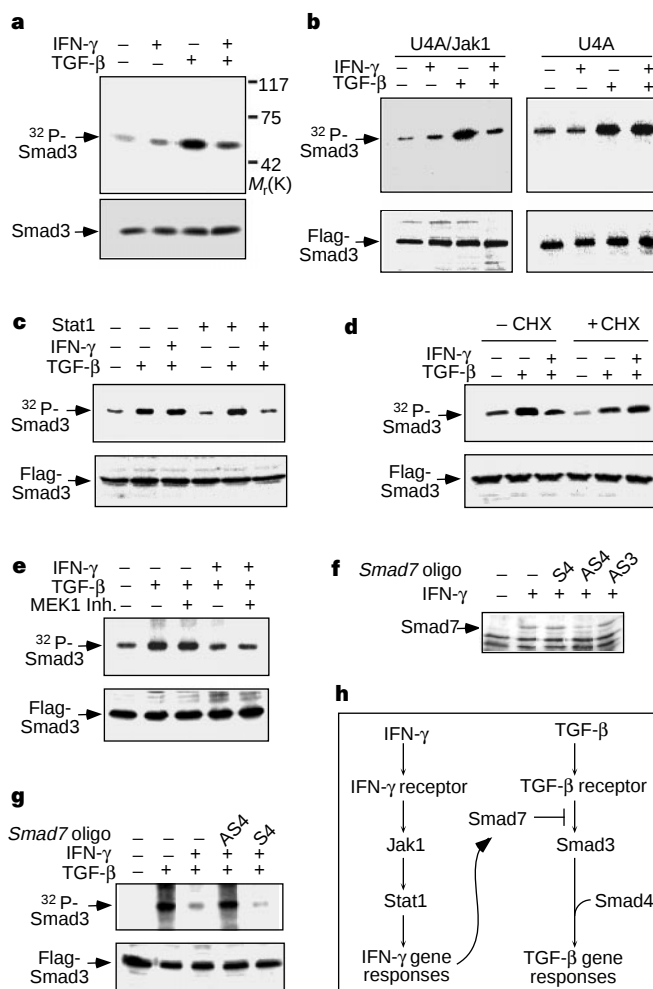


Figure 5 The IFN- γ /Jak1/Stat1 pathway inhibits the TGF- β receptor through *Smad7*. **a**, Phosphorylation of endogenous *Smad3* in [³²P]orthophosphate-labelled U4A/Jak1 cells. Cells were incubated for 1 h with IFN- γ or TGF- β as indicated. *Smad3* levels were monitored by anti-*Smad3* western immunoblotting. **b**, Phosphorylation of Flag-*Smad3* in U4A and U4A/Jak1 cells incubated with cytokines. Flag-*Smad3* levels were monitored by anti-Flag immunoblotting. **c**, Phosphorylation of Flag-*Smad3* in the Stat1-deficient cell line U3A after incubation with cytokines. Where indicated, a Stat1 vector was co-transfected with the Flag-*Smad3* vector. **d**, **e**, Phosphorylation of Flag-*Smad3* in U4A/Jak1 cells incubated with cytokines and **d**, cycloheximide (CHX) or **e**, MEK1 inhibitor added 1 h before cytokine addition. **f**, U4A/Jak1 cells were treated with antisense nucleotides 3 and 4 (AS3, AS4), or sense oligonucleotide 4 (S4), and IFN- γ as indicated. Cell lysates were subjected to anti-*Smad7* immunoblotting. **g**, Effect of *Smad7* antisense oligonucleotide on *Smad3* phosphorylation. Cells were co-transfected with Flag-*Smad3* vector and the indicated *Smad7* oligonucleotides, and then incubated with or without IFN- γ for 45 min before 45 min with or without TGF- β . **h**, A model of the interactions between the STAT and SMAD pathways, suggested by our results. IFN- γ -mediated activation of Jak1 and Stat1 leads to induction of *Smad7*, which inhibits the ability of the TGF- β -receptor complex to interact with, and activate, *Smad3*. As a result, transcriptional responses to TGF- β signalling are inhibited.

oligonucleotide, both of which did not prevent the increase in Smad7, failed to prevent the inhibition of Smad3 phosphorylation by IFN- γ (Fig. 5f, g, and data not shown).

Our results indicate a mechanism for transmodulation of the TGF- β /SMAD signal-transduction pathway by the IFN- γ /STAT pathway: IFN- γ signalling through the Jak1/Stat1 pathway rapidly increases the expression of an anti-SMAD, Smad7, causing the inhibition of TGF β -mediated Smad3 phosphorylation and an attendant loss of TGF- β signalling to the nucleus (Fig. 5h). This mechanism may normally function to balance the activity of these two pathways to provide a suitable level of stimulation under physiological conditions in processes that are under the combined control of TGF- β and IFN- γ ^{2-5,21}. Cytokines that signal through JAK/STAT pathways are generally antagonistic to TGF- β in the regulation of haematopoietic development and immune-cell functions such as inflammation^{1,25}. The mechanism revealed by our results could play a broad part in the integration of STAT and SMAD signalling pathways. □

Method

Cell culture and transfections. All vectors used have been described previously^{14-16,20,26,27}. Transfections were done on exponentially growing cell cultures at 50–70% confluency, which were incubated with MEM medium containing 1 μ g ml⁻¹ DNA and 125 μ g ml⁻¹ DEAE-dextran at 37 °C for 1 h and then shocked for 1 min with 10% dimethyl sulphoxide in PBS at room temperature. Intact or transfected cell cultures plated at 50–60% confluency were treated with agonists as indicated. Luciferase activity was then assayed²⁷. Recombinant human IFN- γ (500 IU ml⁻¹, 2×10^7 IU mg⁻¹; Boehringer), TGF- β 1 (100 pM unless otherwise indicated; R&D Systems), MEK1 inhibitor PD98059 (40 μ M; New England Biolabs) and cycloheximide (20 μ g ml⁻¹; Sigma) were used.

Antibody assays. Cells were lysed by a 10-s sonication in 1 ml TNE buffer (10 mM Tris, pH 8.0, 150 mM NaCl, 1 mM EDTA, 1% NP40) in the presence of protease and phosphatase inhibitors. We used affinity-purified rabbit polyclonal antibodies raised against bacterially expressed glutathione-S-transferase (GST)–Smad3 that recognize Smads 2 and 3 but not Smads 1 or 4. The anti-Flag monoclonal antibody was M2 (Eastman Kodak). Western immunoblotting was done by standard methods and visualized by enhanced chemiluminescence (Amersham). Smad2-specific and Smad3-specific antibodies used in western immunoblotting were raised against synthetic peptides that correspond to residues 81–107 of human Smad2 and residues 192–211 of human Smad3 (Zymed). For immunofluorescence, cells plated on chamber slides at low density for 24 h were stimulated with cytokines as indicated, fixed with methanol/acetone for 2 min, and incubated for 1 h with 10% fetal bovine serum in PBS and for 1 h with affinity-purified anti-Smad2/3 rabbit polyclonal antibodies. Reactions were visualized with biotin-conjugated anti-rabbit IgG (Jackson ImmunoResearch) and fluorescein-isothiocyanate-conjugated streptavidin.

Phosphorylation assays. Cells were labelled with 1 mCi ml⁻¹ [³²P]orthophosphate for 1 h in phosphate-free DMEM medium. After incubation with agonists, cells were lysed in 50 mM Tris, pH 7.4, 150 mM NaCl, 5 mM EDTA, 25 mM NaF, 2 mM phenylmethylsulphonyl fluoride, 1% Triton-X100, in the presence of protease and phosphatase inhibitors. Immunoprecipitated proteins were resolved on SDS–PAGE and subjected to autoradiography. Phosphoamino acid content of [³²P]-labelled Smad3 electrically transferred to immobilized–PVDF membranes was determined by partial hydrolysis in 6 M HCl and two-dimensional separation using the HTLE-7000 system (CBS Scientific).

Antisense oligonucleotides. Phosphorothioate single-stranded oligonucleotides matching regions 77–102 (5'-AGGGCGCAGGGGAGGTGGAGGA-3', oligo 3) and 107–128 (5'-GCTGCGGGGAGAAGGGGCGAC-3', oligo 4) of the human Smad7 complementary DNA sequence (GenBank accession number AF010193) were synthesized in the sense and antisense orientations and used as described²⁸. Briefly, cells were transfected with 12 μ g ml⁻¹ oligonucleotide using lipofectamine in Opti-MEM1 media (Gibco BRL) for 6 h, followed by incubation in media with oligonucleotide but without lipofectamine for 12 h. Cells were then incubated in serum-free media for 4 h ($\pm$ oligonucleotide), and

treated with IFN- γ for 45 min followed by TGF- β for another 45 min, as indicated.

Other assays. Northern assays of poly(A)⁺ RNA samples were done using cDNA probes for human Smad6 (ref. 29), Smad7 (ref. 12), FKBP12 and glyceraldehyde-3-phosphate dehydrogenase. For assays of receptor–Smad interactions, cells were transiently transfected with T β RI and T β RII vectors (for Smad7–receptor-interaction assays) or with Flag–Smad3, T β RI and T β RII vectors (for assays of Flag–Smad3–receptor interactions). Cells were incubated with IFN- γ , then crosslinked to cell-bound ¹²⁵I-labelled TGF- β in the cold³⁰ and immunoprecipitated with anti-Flag antibody. Labelled receptors in the immunoprecipitates and total cell lysates were visualized by SDS–PAGE and autoradiography.

Received 21 December 1998; accepted 19 January 1999.

- Letterio, J. L. & Roberts, A. B. Regulation of immune responses by TGF- β . *Annu. Rev. Immunol.* **16**, 137–161 (1998).
- Czarniecki, C. W., Chiu, H. H., Wong, C. H., McCabe, S. M. & Palladino, M. A. Transforming growth factor- β 1 modulates the expression of class II histocompatibility antigens on human cells. *J. Immunol.* **140**, 4217–4223 (1988).
- Bauvois, B., Rouillard, D., Sanceau, J. & Wietzerbin, J. IFN- γ and transforming growth factor- β 1 differently regulate fibronectin and laminin receptors of human differentiating monocytic cells. *J. Immunol.* **148**, 3912–3919 (1992).
- Schmitt, E. *et al.* T helper type I development of naive CD4⁺ T cells requires the coordinate action of interleukin-12 and interferon- γ and is inhibited by transforming growth factor- β . *Eur. J. Immunol.* **24**, 793–798 (1994).
- Xiao, B. G., Zhang, G. X., Ma, C. G. & Link, H. Transforming growth factor- β 1 (TGF- β 1)-mediated inhibition of glial cell proliferation and down-regulation of intercellular adhesion molecule-1 (ICAM-1) are interrupted by interferon- γ (IFN- γ). *Clin. Exp. Immunol.* **103**, 475–481 (1996).
- Stark, G. R., Kerr, I. M., Williams, B. R., Silverman, R. H. & Schreiber, R. D. How cells respond to interferons. *Annu. Rev. Biochem.* **67**, 227–264 (1998).
- Massagué, J. TGF β signal transduction. *Annu. Rev. Biochem.* **67**, 753–791 (1998).
- Heldin, C.-H., Miyazono, K. & ten Dijke, P. TGF- β signalling from cell membrane to nucleus through SMAD proteins. *Nature* **390**, 465–471 (1997).
- Schindler, C. & Darnell, J. E. Transcriptional responses to polypeptide ligands: the Jak-Stat pathway. *Annu. Rev. Biochem.* **64**, 621–651 (1995).
- Ilhe, J. N. Cytokine receptor signalling. *Nature* **377**, 591–594 (1995).
- Hayashi, H. *et al.* The MAD-related protein Smad7 associated with the TGF β receptor and functions as an antagonist of TGF β signalling. *Cell* **89**, 1165–1173 (1997).
- Nakao, A. *et al.* Identification of Smad7, a TGF β -inducible antagonist of TGF- β signalling. *Nature* **389**, 631–635 (1997).
- Muller, M. *et al.* The protein tyrosine kinase Jak1 complements defects in interferon- α / β and - γ signal transduction. *Nature* **366**, 129–135 (1993).
- Cacamo, J. *et al.* Type I receptors specify growth inhibitory and transcriptional responses to TGF- β and activin. *Mol. Cell Biol.* **14**, 3810–3821 (1994).
- Chen, X., Rubock, M. J. & Whitman, M. A transcriptional partner of MAD proteins in TGF- β signalling. *Nature* **383**, 691–696 (1996).
- Liu, F., Poupponnot, C. & Massagué, J. Dual role of the Smad4/DPC4 tumor suppressor in TGF β -inducible transcriptional responses. *Genes Dev.* **11**, 3157–3167 (1997).
- Kretschmar, M., Doody, J. & Massagué, J. Opposing BMP and EGF signalling pathway converge on the TGF β family mediator Smad1. *Nature* **389**, 618–622 (1997).
- Sakatsume, M. *et al.* Interferon γ activation of Raf-1 is Jak1-dependent and p21ras-independent. *J. Biol. Chem.* **273**, 3021–3026 (1998).
- Dudley, D. T., Pang, L., Decker, S. J., Bridges, A. J. & Saltiel, A. R. A synthetic inhibitor of the mitogen-activated protein kinase cascade. *Proc. Natl Acad. Sci. USA* **92**, 7686–7689 (1995).
- Muller, M. *et al.* Complementation of a mutant cell line: central role of the 91 kDa polypeptide of the ISGF3 in the interferon- α and - γ signal transduction pathways. *EMBO J.* **12**, 4221–4228 (1993).
- Spittler, A. *et al.* Effects of 1 α ,25-dihydroxyvitamin D3 and cytokines on the expression of MHC antigens, complement receptors and other antigens on human blood monocytes and U937 cells: role in cells differentiation, activation and phagocytosis. *Immunology* **90**, 286–293 (1997).
- Lee, K., Tanaka, M., Hatanaka, M. & Kuze, F. Reciprocal effects of epidermal growth factor and transforming growth factor and transforming growth factor β on the anchorage-dependent and -independent growth of A431 epidermoid carcinoma cells. *Exp. Cell Res.* **173**, 156–162 (1987).
- Kumar, R. & Mendelsohn, J. Growth regulation of A431 cells. Modulation of expression of transforming growth factor- α mRNA and 2',5'-oligoadenylate synthetase activity. *J. Biol. Chem.* **265**, 4578–4582 (1990).
- Souchelnytskyi, S. *et al.* Physical and functional interaction of murine and *Xenopus* Smad7 with bone morphogenetic protein receptors and transforming growth factor- β receptors. *J. Biol. Chem.* **273**, 25364–25370 (1998).
- Ohta, M., Greenberger, J. S., Anklesaria, P., Bassols, A. & Massagué, J. Two forms of transforming growth factor- β distinguished by multipotential hematopoietic progenitor cells. *Nature* **329**, 539–541 (1987).
- Zhang, Y., Feng, X.-H., Wu, R.-Y. & Derynck, R. Receptor-associated Mad homologues synergize as effectors of the TGF β response. *Nature* **383**, 168–172 (1996).
- Lagna, G., Hata, A., Hemmati-Brivanlou, A. & Massagué, J. Partnership between DPC4 and SMAD proteins in TGF- β signalling pathways. *Nature* **383**, 832–836 (1996).
- Tu, G. C., Cao, Q.-N., Zhou, F. & Israel, Y. Tetranucleotide GGGA motif in primary RNA transcripts. Novel target for antisense design. *J. Biol. Chem.* **273**, 25125–25131 (1998).
- Hata, A., Lagna, G., Massagué, J. & Hemmati-Brivanlou, A. Smad6 inhibits BMP/Smad1 signaling by specifically competing with the Smad4 tumor suppressor. *Genes Dev.* **12**, 186–197 (1998).
- Massagué, J. Identification of receptors of type β transforming growth factor. *Methods Enzymol.* **146**, 174–195 (1987).

Acknowledgements. We thank G. Stark, I. Kerr, R. Flavell, C. H. Heldin, P. Ten Dijke, R. Derynck and M. Whitman for cell lines and vectors; A. Hata, F. Liu, A. Nieporent and J. Ilhe for useful insights; and C. Zhang for technical assistance. This work was supported by an NIH grant to J.M. and a Cancer Center grant to the Memorial Sloan-Kettering Cancer Center. L.U. is a postdoctoral fellow of the International Human Frontier Organization Program. J.M. is an investigator of the Howard Hughes Medical Institute.

Correspondence and requests for materials should be addressed to J.M. (e-mail: j-massague@ski.mskcc.org).

Structure of a DNA-bound Ultrabithorax–Extradenticle homeodomain complex

Jonathan M. Passner*, Hyung Don Ryoo†, Leyi Shen*, Richard S. Mann† & Aneel K. Aggarwal*

* Structural Biology Program, Department of Physiology and Biophysics, Box 1677, Mount Sinai School of Medicine, 1425 Madison Avenue, New York, New York 10029, USA

† Department of Biochemistry and Molecular Biophysics, Columbia University, New York, New York 10032, USA

During the development of multicellular organisms, gene expression must be tightly regulated, both spatially and temporally. One set of transcription factors that are important in animal development is encoded by the homeotic (Hox) genes, which govern the choice between alternative developmental pathways along the anterior–posterior axis^{1,2}. Hox proteins, such as *Drosophila* Ultrabithorax, have low DNA-binding specificity by themselves but gain affinity and specificity when they bind together with the homeoprotein Extradenticle (or Pbx1 in mammals)^{3,4}. To understand the structural basis of Hox–Extradenticle pairing, we determine here the crystal structure of an Ultrabithorax–Extradenticle–DNA complex at 2.4 Å resolution, using the minimal polypeptides that form a cooperative heterodimer. The Ultrabithorax and Extradenticle homeodomains bind opposite faces of the DNA, with their DNA-recognition helices almost touching

each other. However, most of the cooperative interactions arise from the YPWM amino-acid motif of Ultrabithorax—located amino-terminally to its homeodomain—which forms a reverse turn and inserts into a hydrophobic pocket on the Extradenticle homeodomain surface. Together, these protein–DNA and protein–protein interactions define the general principles by which homeotic proteins interact with Extradenticle (or Pbx1) to affect development along the anterior–posterior axis of animals.

There are eight *Hox* genes in *Drosophila*, all of which encode transcription factors that bind DNA through their homeodomains^{2,5} (Fig. 1). The homeodomain structure consists of a tri- α -helical core and an amino-terminal arm that becomes partially ordered in the presence of DNA^{5,6}. The Extradenticle (Exd) homeodomain is one of the most divergent from the canonical homeodomain, with two unusual DNA-recognition residues and an extra three residues in the loop between helices $\alpha 1$ and $\alpha 2$. Exd is similar to the human proto-oncoprotein Pbx1, which binds DNA cooperatively with the mammalian Hox proteins^{3,4}, as Exd does with *Drosophila* Hox proteins. Most Hox proteins are characterized by the presence of a YPWM amino-acid motif at a variable distance upstream of their homeodomain; this motif is necessary for optimal Hox–Exd cooperativity (Fig. 1). In solving the Hox–Exd–DNA structure, we asked the following questions. How does an unusual homeodomain, like that of Exd, bind to DNA? What are the molecular mechanisms underlying Hox–Exd cooperativity? In particular, what are the roles of the Hox YPWM motif and the extra long loop between helices $\alpha 1$ and $\alpha 2$ in Exd? How similar or different is the Hox–Exd heterodimer from other homeodomain structures?

Exd and Ultrabithorax (Ubx) bind DNA four base pairs (bp) apart, with their homeodomains aligned in a head-to-tail orienta-

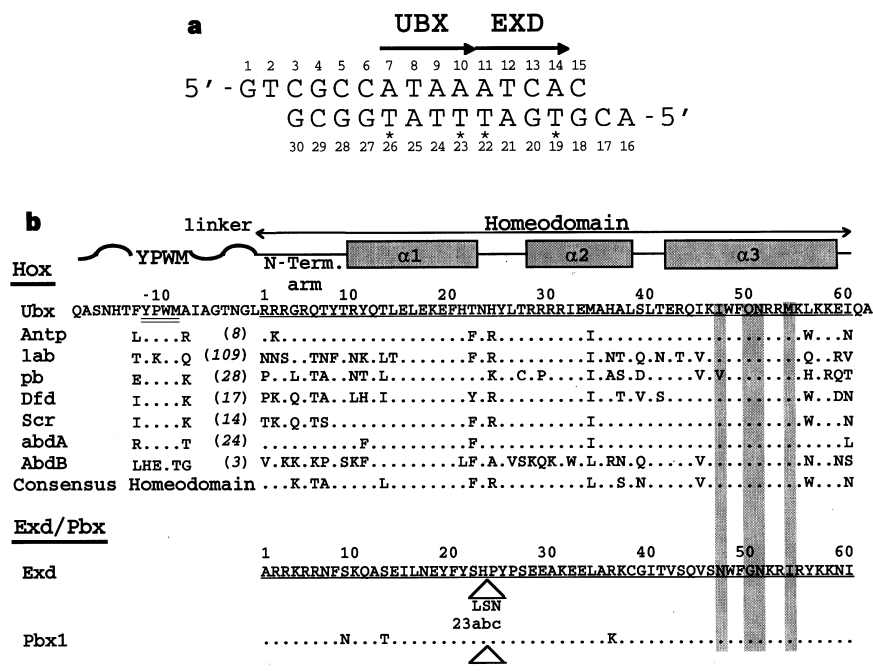


Figure 1 DNA and protein sequences. **a**, Sequence of the 15-bp DNA fragment used to crystallize the ternary complex. The arrows show the DNA-core-binding sequences of Ubx (ATAA) and Exd (ATCA) and highlight the head-to-tail arrangement of the two homeodomains. Asterisks denote thymine residues substituted by iodouracils in the two iodinated derivatives (Table 1). This oligonucleotide forms cooperative complexes with multiple Hox–Exd heterodimers (data not shown). **b**, The sequence of the Ubx construct used here (in the wild-type protein residue 39 is a cysteine) is aligned with the seven other *Drosophila* Hox proteins. The homeodomain is numbered from 1 to 60, and residues N-terminal to the homeodomain that contain the YPWM motif are

numbered from –1 to –19. The distances between the YPWM motifs and the homeodomains are shown in parentheses. There is no consensus YPWM motif in AbdB but a conserved tryptophan is required to form cooperative heterodimers with Exd¹³. Shown above the homeodomain sequences are the relative locations of the α -helices. Bottom, the sequence of the Exd construct used here, aligned with the sequence of the human oncoprotein Pbx1. The three-amino-acid insert in the Exd/Pbx1 homeodomains is labelled with residues 23a, 23b and 23c. The residues that normally contact DNA bases in the major groove in homeodomains are shaded.

tion (Fig. 2), consistent with DNA-binding and mutagenesis studies^{7–9}. The Exd homeodomain is tailored for protein–protein interactions. The extra three residues between helices $\alpha 1$ and $\alpha 2$ of Exd bulge out to form part of the hydrophobic pocket that binds the Ubx YPWM motif (Fig. 3a). The Exd–DNA interactions are unusual for homeodomain–DNA interactions (Fig. 3b, c). The traditional view of such interactions involves an ATTA core in the DNA; the first two base pairs of this core usually contact the homeodomain in the DNA's major groove, and the last two base pairs contact the homeodomain in the minor groove^{5,6}. The matching sequence for Exd in our structure is ATCA. The usual hydrophobic contacts (from Val 47 or Ile 47 of the homeodomain) to the thymine methyl group of the first base pair of the core sequence are missing in Exd, and are replaced by direct and water-mediated hydrogen bonds between Asn 47 and the DNA backbone. (An asparagine at position 47 is unusual and found in the Exd, Pbx and Mat $\alpha 2$ homeodomains.) Similarly, the normal contacts to bases preceding the ATCA core sequence are absent because a glycine, instead of a more typical glutamine, lysine or serine residue, is present at position 50; the contacts are replaced by a water-mediated interaction between the Gly 50 carbonyl and the adenine preceding the ATCA core. However, the lack of these specific contacts at the 5' end of the core sequence is made up to some extent at the 3' end, by Arg 55 of Exd making bidentate hydrogen bonds with the guanine of the third base pair of the ATCA sequence. A similar interaction is made by Mata1 in the Mat $\alpha 2$ –Mata1–DNA structure¹⁰, but it is otherwise uncommon for homeodomains. An interaction that Exd shares with all homeodomain structures, including Ubx here, is the bidentate hydrogen bonds between Asn 51 and the adenine of the second base pair of ATCA.

There are several other unexpected features to Exd–DNA interactions. First, whereas Tyr 25, Arg 53 and Lys 57 interact with the DNA phosphate backbone, the hydrophobic portions of these residues form part of the YPWM-motif binding pocket (see later). In addition, the N terminus of helix $\alpha 2$ (residues 28–32), which is close to the negatively charged DNA backbone, is very basic in the consensus homeodomain sequence. However, in Exd many of these residues are replaced by glutamic acid. The net charge difference in this region between Exd and a consensus homeodomain is –7, which, according to electrostatic considerations, should considerably weaken the affinity of Exd for DNA. This is borne out in solution, where Exd binds DNA poorly on its own^{7,11}. The Exd

homeodomain thus appears to be tailored for cooperative interactions—its hydrophobic pocket, its modified recognition site (ATCA), and its weakened DNA-binding ability are all features that contribute to its requirement for a Hox protein for effective DNA binding.

The Ubx structure can be divided into its homeodomain (residues 1–60) and the N-terminal extension, which carries the YPWM motif (residues –1 to –19). Given the unusual features of the Exd–DNA interactions and the close proximity of the Exd- and Ubx-binding sites in DNA, it was possible that the Ubx homeodomain was altered in conformation. However, unlike Exd, the Ubx homeodomain appears remarkably similar to previously characterized homeodomains. A superposition of the Ubx homeodomain with the X-ray-derived Antp homeodomains¹² gives an average root-mean-square deviation (r.m.s.d.) of only 0.41 Å (for C α atoms of residues 8–56). This is comparable to the 0.2 Å r.m.s.d. between independently determined Antp homeodomains¹², suggesting at most only a mild conformational change when Hox homeodomains bind DNA by themselves or as heterodimers with Exd. This invariance in structure extends even to the Ubx N-terminal arm, the first four residues of which are disordered as in the Antp homeodomain. The conformation of the DNA, however, changes markedly upon heterodimerization with a significant broadening of its minor groove. For instance, the r.m.s.d. between phosphates of the DNAs comprising the Ubx- and Antp-recognition sites (CCAT{T/A}A) is 1.18 Å, in comparison to an r.m.s.d. of 0.49 Å between the independently determined DNAs in the Antp structure¹². Ubx interacts with DNA only through its homeodomain, contacting the DNA backbone and the base pairs in the major and the minor grooves of the core DNA sequence (ATAA) in a manner typical of most homeodomains^{5,6} (Fig. 3b, d).

Table 1 Crystallographic analysis

	Native 1	Native 2	lodo 1*	lodo 2*	SeMet
Resolution (Å)	20–2.4	25–2.8	25–2.8	25–3.1	20–3.5
Number of reflections					
Measured	110,271	49,177	39,830	41,189	27,379
Unique	10,324	6,537	6,049	4,597	3,514
Data coverage (%)	98.3 (100)	99.9 (100.0)	93.9 (100.0)	95.0 (96.6)	97.6 (99.7)
R_{sym} (%)†	7.4 (18.5)	5.8 (29.7)	7.1 (25.5)	8.4 (21.9)	5.1 (6.2)
Phasing statistics (20–3.5 Å)					
Phasing power (isomorphous)‡			1.50	1.43	
R_{cullis} (anomalous)§			0.65	0.74	
Mean figure of merit		0.66			
Refinement statistics (8–2.4 Å)					
Reflections, $F > 2\sigma(F)$	9,490				
R-factor (%)	22.4				
R-free (%)	30.4				
Non-hydrogen atoms	1,747				
r.m.s. bond lengths (Å)	0.005				
r.m.s. bond angles (°)	0.95				

Values shown in parentheses are for the highest resolution bin (2.49–2.4 Å for native 1, 2.9–2.8 Å for native 2 and lodo 1, 3.21–3.1 Å for lodo 2 and 3.62–3.5 Å for SeMet). r.m.s., root mean square.

* The derivatives lodo 1 and lodo 2 consist of ioduracil replacing thymines at bases 19 and 23 and at bases 22 and 26, respectively.

† $R_{\text{sym}} = \sum |I_{\text{obs}} - \langle I \rangle| / \sum I$, calculated for all data.

‡ Phasing power (isomorphous) = $\langle |F_{\text{calc}}| \rangle / \langle |F_{\text{calc}}| \rangle$, where $\langle |F_{\text{calc}}| \rangle$ is the mean calculated amplitude for the heavy-atom model and ϵ is the lack of closure error; acentric reflections only.

§ $R_{\text{cullis}} = \text{r.m.s. lack of closure error/average anomalous difference}$.

|| Calculated with 5% of the reflection data excluded from refinement.

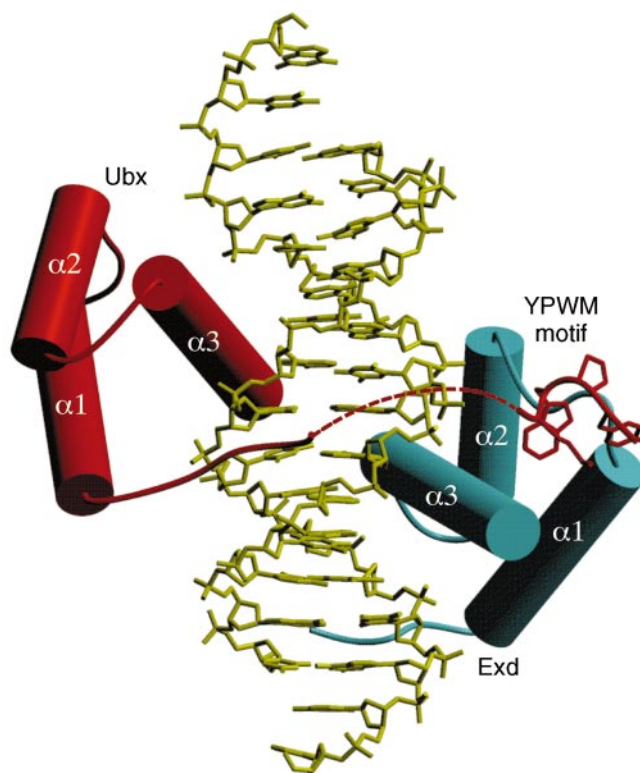


Figure 2 An overview of the Ubx–Exd–DNA ternary complex. The diagram shows the Ubx (red) and Exd (cyan) homeodomains binding in a tandem manner on opposite faces of the DNA (yellow). The dashed red line represents the disordered linker between the Ubx homeodomain and its YPWM motif. The YPWM motif reaches into a hydrophobic pocket on the surface of the Exd homeodomain. Figs 2, 3 and 4b were generated with SETOR²⁹.

The Ubx YPWM motif is ordered from residues -13 to -8, which comprise the FYPWMA segment of the polypeptide. The intervening residues connecting the YPWM motif to the homeodomain are disordered, implying flexibility within this linker region, in agreement with the variable distance between the YPWM motif and the homeodomain in different Hox proteins (Fig. 1b). The YPWM sequence folds into a classical type I reverse turn, with a hydrogen bond between the first (tyrosine) and the fourth (methionine) residues of the turn (Fig. 4a). We found the YPWM sequence in two unrelated proteins in the Brookhaven Protein Database (acces-

sion numbers 2MTA and 1KB5), and in each the tetrapeptide formed a reverse turn as in our structure. Thus, the identity of amino acids within the YPWM sequence appears to be dictated both by their interactions with the Exd homeodomain and by their propensity to form a reverse turn.

Ubx and Exd cooperate through a combination of protein-protein and protein-DNA interactions. The protein-protein interactions are spread over several segments of the Ubx and Exd polypeptides (Fig. 4). Visually, the most striking interactions arise from the insertion of the YPWM motif of Ubx into a hydrophobic

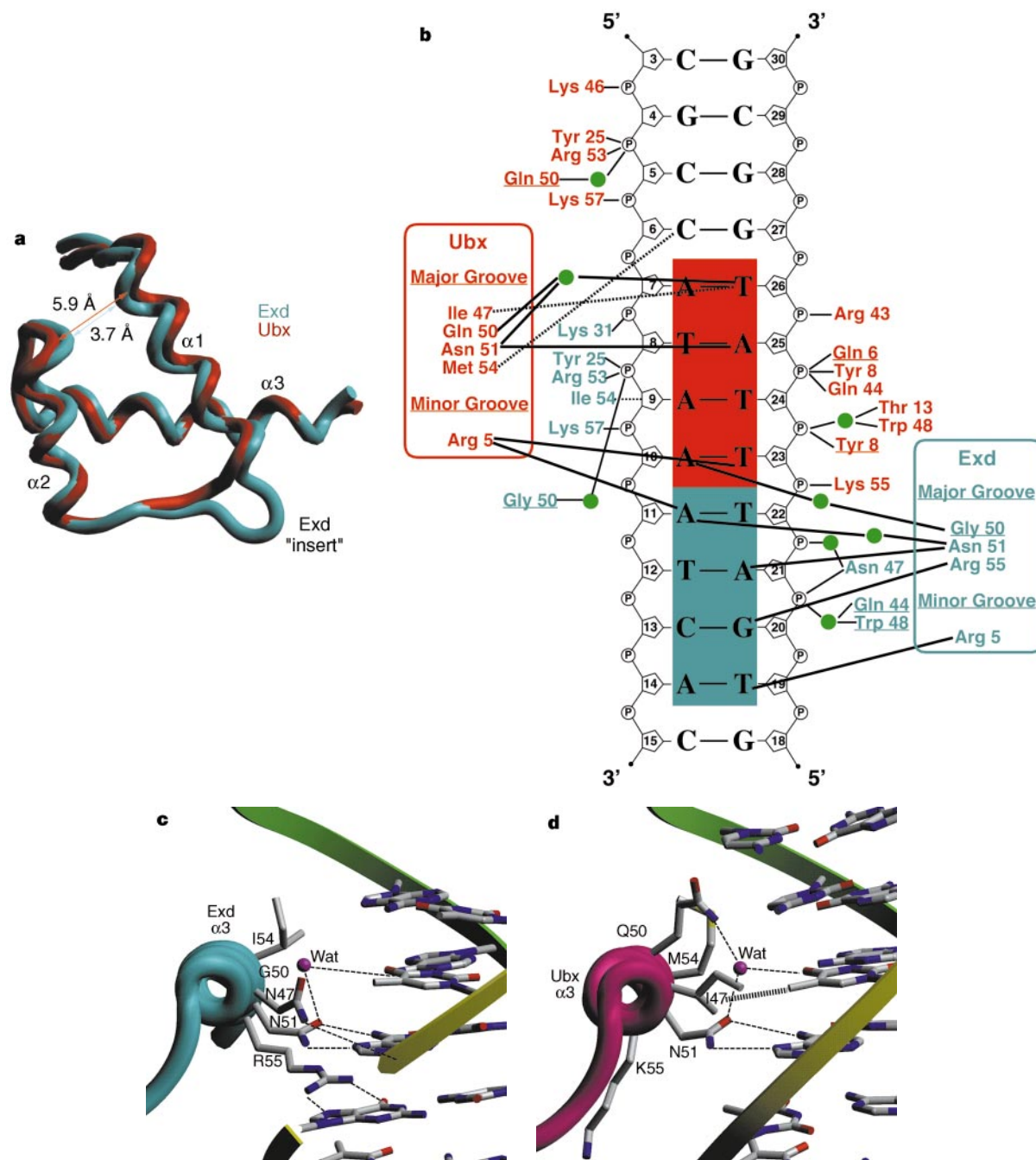


Figure 3 Exd-Ubx structure and DNA interactions. **a**, A superposition of the Exd (cyan) and Ubx (red) homeodomains (r.m.s.d. of 0.92 Å), done using the C α atoms of residues 8-56 (excluding the Exd three-amino-acid insert). **b**, A summary of the protein-DNA contacts. The Ubx residues and its core binding site in DNA (ATAA) are shown in red, and the Exd residues and its core DNA site (ATCA) are shown in cyan. Hydrogen bonds are represented by solid lines and nonpolar interactions by dotted lines. Interactions involving the protein main chain are underlined. Water molecules are shown as green circles. **c**, A view along the Exd DNA-

recognition helix, showing residues Asn 47, Asn 51, Ile 54 and Arg 55 interacting with bases in the DNA major groove. Also shown is Gly 50; the corresponding residue in Ubx (Gln 50) makes base contacts. Hydrogen bonds are shown as dashed lines. **d**, A view along the Ubx DNA-recognition helix, showing residues Ile 47, Gln 50, Asn 51 and Met 54 interacting with bases in the DNA major groove. Also shown is residue Lys 55; the corresponding residue in Exd (Arg 55) makes base contacts. Van der Waals contacts between Ile 47 and a thymine base are represented by a thick dashed line.

pocket on the Exd surface (Fig. 4a). The Exd hydrophobic pocket is lined by residues from the helix $\alpha 1$ N terminus (Phe 20), the extra long loop between helices $\alpha 1$ and $\alpha 2$ (Leu 23a, Ser 23b, Asn 23c (see legend to Fig. 1 for an explanation of this nomenclature), Pro 24 and Tyr 25), and the DNA-recognition helix (helix $\alpha 3$) carboxy terminus (Arg 53, Tyr 56 and Lys 57) (Fig. 4b). A canonical homeodomain (such as that of Ubx) lacks an equivalent hydrophobic pocket.

The binding of the Ubx YPWM motif within the Exd hydrophobic pocket leads to a total loss of $\sim 570 \text{ \AA}^2$ in solvent-accessible surface area from both proteins. The tryptophan residue of the

motif plays a dominant part in the binding, mediating many of the hydrophobic contacts within the pocket and making a hydrogen bond to the main-chain carbonyl of Leu 23a. The importance of the tryptophan residue is underscored by its absolute conservation in the more than 70 YPWM motifs now known. A conserved tryptophan has also been implicated in the cooperative interactions between AbdB homologues and Pbx1a, despite the absence of a consensus YPWM motif in these Hox proteins¹³. The tyrosine, proline and methionine residues of the Ubx YPWM motif make less extensive interactions with Exd, but their importance also lies in

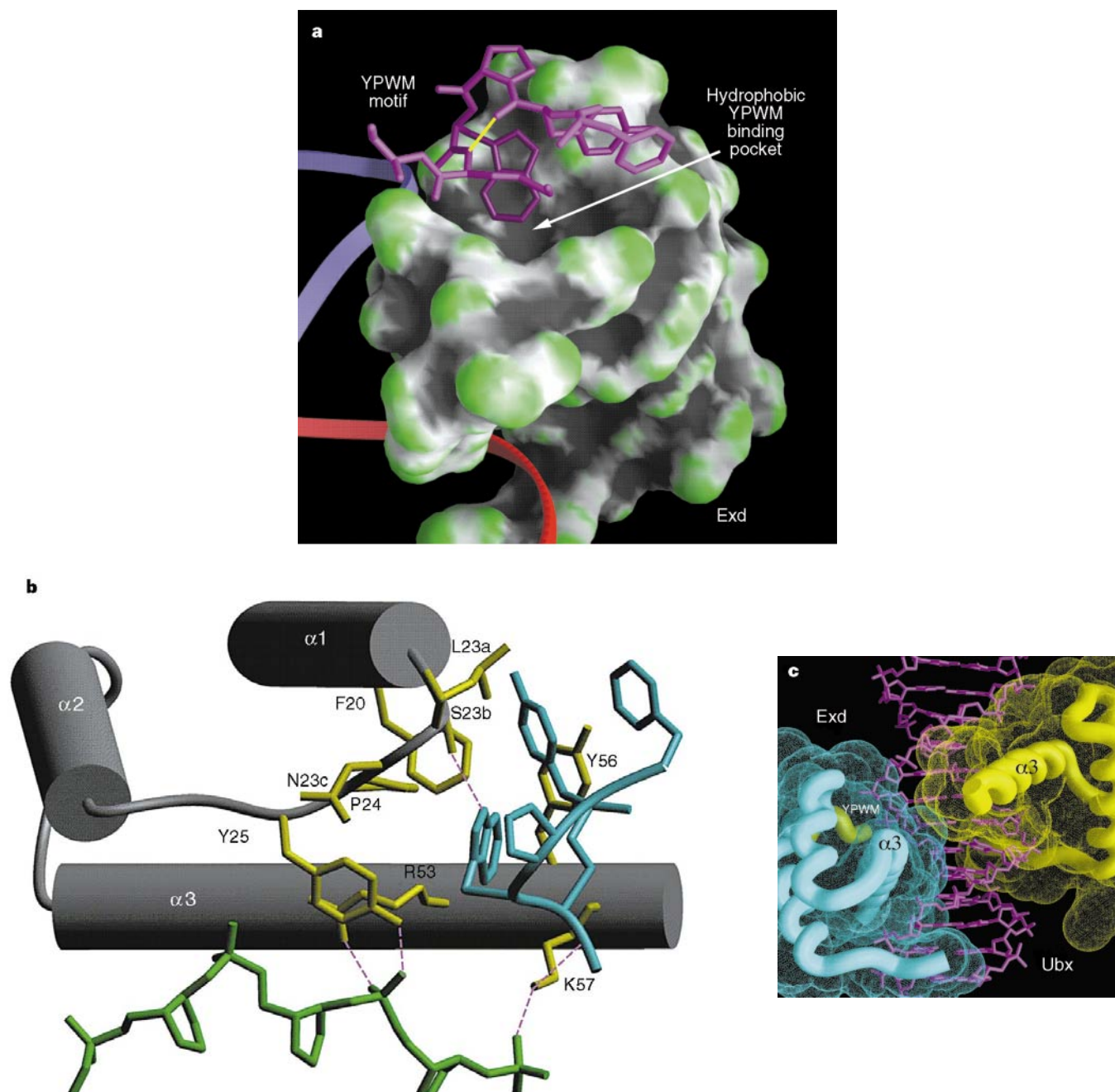


Figure 4 Protein-protein interactions. **a**, The YPWM motif of Ubx reaches into a hydrophobic pocket on the Exd homeodomain surface. The YPWM motif (magenta) forms a type 1 reverse turn, with a hydrogen bond (yellow) between the first and fourth residues of the motif. The surface of Exd is coloured according to its curvature, with convex patches shaded green and concave patches shaded grey. The DNA backbone (purple and red) is included as a reference. The overall orientation is similar to that in Fig. 2. **a, c**, Were made with GRASP³⁰. **b**, A detailed view of the interactions between the Ubx YPWM motif (cyan) and the residues

within the Ubx-binding pocket (yellow) of Exd (grey). Several of the residues (Tyr 25, Arg 53 and Lys 57) that line the Exd hydrophobic pocket also contact the DNA backbone (green). Hydrogen bonds are shown by dashed lines. **c**, A 'back' view of the ternary complex, obtained by rotating 180° about the DNA-helical axis from the orientation shown in Fig. 2. The figure shows the close approach of the Ubx (yellow) and Exd (cyan) homeodomains, with their DNA-recognition helices almost touching each other. The transparent surfaces represent the water-exclusion surfaces of the two proteins.

helping to buttress the tryptophan residue in the hydrophobic pocket through stacking interactions. Similarly, we observe the involvement of residues outside the motif (phenylalanine and alanine of the sequence FYPWMA); in particular, a hydrogen bond is formed between the alanine main-chain amide and Exd residue Lys 57. Remarkably, several of the residues that line the Exd hydrophobic pocket also contact the DNA backbone. As noted above, the nonpolar portions of Tyr 25, Arg 53 and Lys 57 side chains form part of the Exd hydrophobic pocket, whereas their polar/charged groups make hydrogen bonds with the DNA backbone (Fig. 4b). This may explain why the addition of a 12-amino-acid peptide containing the Hox A5 YPWM motif enhances DNA binding by Pbx1 (ref. 7). We surmise that the peptide increases DNA binding by Pbx1 at least in part by reducing the configurational entropy of the Tyr 25, Arg 53 and Lys 57 side chains in their interactions with DNA. Apparently, the YPWM motif of Hox proteins confers cooperativity both through protein–protein interactions and by helping to configure the Exd/Pbx1 homeodomain for optimal DNA contacts.

An interaction between Hox and Exd/Pbx1 proteins that is independent of the YPWM motif has also been reported^{11,14,15}. Because the Ubx and Exd homeodomains are juxtaposed closely, their DNA-recognition helices almost touch each other (Fig. 4c). The Ubx recognition-helix C terminus is located within 9 Å of the Exd homeodomain recognition-helix N terminus, with Lys 58 of Ubx in a position to form a hydrogen bond with Ser 43 of Exd. However, Lys 58 is partially disordered in our crystal structure, but it may rotate in the direction of Ser 43 in solution. Nonetheless, even without this specific interaction there is a significant loss in solvent-accessible surface area ($\sim 50 \text{ Å}^2$) because of the proximity of the two DNA-recognition helices, which may account for some of the weak cooperativity observed without the YPWM motif. Residues located C-terminally to Hox homeodomains have also been implicated in YPWM-independent cooperativity¹¹ and in Hox specificity⁴; we imagine that these residues may interact with Exd. The proximity of the two homeodomains may also explain an intriguing Glu 28 → Arg mutation in Pbx1 that enhances DNA binding by monomeric Pbx1 but abolishes cooperative binding⁷. This mutation is similar to the naturally occurring hypomorphic *exd* mutation, *exd*^{EM52}, in which this glutamate is changed to a lysine¹⁶. In our structure, Glu 28 lies at the N terminus of Exd helix $\alpha 2$, away from the YPWM-binding pocket but close to the negatively charged DNA backbone. The substitution of Glu 28 with a basic arginine or lysine may alter the way in which Pbx1 or Exd binds the DNA, resulting in possible steric clashes with the adjacently bound Hox homeodomain.

The close proximity of the two homeodomains produces highly overlapping protein–DNA contacts. Exd makes more than half of its contacts with the DNA backbone within the Ubx DNA-recognition site (Fig. 3b). Thus, Exd binding may preconfigure the DNA for Ubx recruitment by affecting both the structure and the flexibility of the DNA around Ubx's DNA-recognition site. A similar mechanism of cooperativity has been proposed for the transcription factor Oct-1, which contains the bipartite POU DNA-binding domain; this domain is composed of a homeodomain and a POU-specific domain joined by a flexible linker. The two subdomains bind to opposite faces of the DNA in a cooperative manner and make overlapping contacts with the DNA backbone, even after the removal of the connecting linker and without any other protein–protein contacts¹⁷. Thus both protein–protein and protein–DNA interactions form the basis of Ubx–Exd partnership. The major cooperative interactions result from the extension of the YPWM motif of Ubx into the hydrophobic pocket on the Exd surface. However, a significant component of the cooperativity can be attributed to the proximity of the two homeodomains on the DNA.

Exd has the ability to form heterodimers preferentially with the one Hox protein over another depending on subtle variations in the

two variable base pairs (N) within the generalized heterodimerization site on the DNA, ATNNATCA (ref. 3). At the same time, the first few residues of the Hox homeodomain N-terminal arm are also important for the functional specificities of Hox proteins *in vivo*⁴ and contribute to the different Hox–Exd cooperativities^{9,18}. These facts have led to the hypothesis that Exd enhances Hox specificity, at least in part, by altering the way in which the Hox N-terminal arm interacts with these base pairs^{9,18}. However, in the structure, no specific interactions are observed between the N-terminal arm and the variable bases. Hence, it is possible that the *in vivo* interaction is helped by another protein or that the identity of the base pairs is dictated by indirect read-out of the phosphate backbone, like the read-out achieved by the phage 434 or trp repressors^{19,20}. To fully understand the subtleties of Hox selectivity, a detailed comparison of several Hox–Exd structures will be required.

In comparing the Ubx–Exd structure to the Mat α 2–Mata1 and Mat α 2–MCM1 structures^{10,21}, several common and unique mechanisms governing homeodomain–cofactor interactions emerge. The first common theme is the use of regions close to the homeodomain for protein–protein interactions. Analogous to the YPWM motif of Hox proteins, Mat α 2 uses regions on both sides of its homeodomain for interactions with MCM1 and Mata1. In all three complexes, hydrophobic interactions play a central part, with transiently ordered regions filling hydrophobic pockets on the cofactor surfaces. The buried solvent-accessible surface areas are 570 Å^2 for Ubx–Exd, 750 Å^2 for Mat α 2–Mata1 and $1,200 \text{ Å}^2$ for Mat α 2–MCM1. All three values are below those observed for most oligomeric proteins²², consistent with homeoproteins existing as monomers in solution and only forming heterodimers on DNA.

The structures also illustrate the versatility of the DNA in accommodating different heterodimers, bending to allow protein–protein interactions when the two proteins bind to the same face of the DNA, as in the Mat α 2 complexes, or remaining straight when the two proteins are on opposite faces of the DNA, as here. This variation is also reflected in the dramatic difference in the spacing between the two homeodomain-binding sites, which are 12 base pairs apart in the Mat α 2–Mata1 complex but only 4 base pairs apart in the Hox–Exd complex. Correspondingly, in the Mat α 2–Mata1 complex there are no homeodomain–homeodomain contacts, whereas in the Hox–Exd structure it is possible that such contacts exist. Overall, a comparison between the Mat α 2–Mata1 and Hox–Exd complexes reveals an underlying flexibility in the way in which homeodomain proteins form multimeric complexes. Perhaps this flexibility is the reason why this DNA-binding motif is used in so many different ways among eukaryotes. □

Methods

Data collection. The *Drosophila* Ubx (residues 233–313 of Ubx isoform IVa) and the Exd (residues 238–300) constructs were expressed, purified and co-crystallized with a 15-bp DNA oligonucleotide in 0.6 M ammonium phosphate (J.M.P., H.D.R., L.S., R.S.M. and A.K.A., manuscript in preparation). When cooled to -160°C , the co-crystals diffracted to 2.8 Å on a rotating anode X-ray source (native 2) and to 2.4 Å at the Cornell High Energy Synchrotron Source (CHESS) on beamline F1 (native 1). The crystals belong to space group $P4_32_12$ with unit-cell dimensions of $a = b = 69.4 \text{ Å}$, $c = 100.7 \text{ Å}$, and contain a single ternary complex per asymmetric unit. Each of the Iodo derivatives was prepared by substituting iodouracils for two thymines in the DNA fragment (Fig. 1a). All data sets were processed using the HKL programs²³.

Structure determination and refinement. The iodine sites were located from both isomorphous and anomalous-difference Patterson maps, and their positions refined with MLPHARE²⁴. The resulting multiple isomorphous replacement including anomalous scattering (MIRAS) phases were used to calculate a 3.5 Å map that was continuous and showed clear electron density for the DNA, the Ubx homeodomain and the Exd-homeodomain helices. An idealized B-DNA fragment was fit into the density and was consistent with the iodine positions. A polyaniline homeodomain without its N-terminal arm, derived from the high-resolution even-skipped structure²⁵, was placed into the

Ubx-homeodomain electron density, and polyaniline helices were placed into the Exd-homeodomain density. The MIRAS map was further improved by cycles of solvent flattening with program DM²⁴. Using these and $F_o - F_c$ and $2F_o - F_c$ maps, interspersed with positional and individual B-factor refinement using X-PLOR²⁶, the model was rebuilt, side chains were added, and the Exd loops and N-terminal arms were built with the program O (ref. 27). The first four residues of Exd, the residues from -7 to 4 of Ubx and the first 6 residues of Ubx were disordered. There was clear density for the YPWM motif, but it was not readily interpretable for residues other than the tryptophan side chain. To obtain a more interpretable map, we refined and improved the MIRAS phases by solvent flattening and extended them to 2.8 Å using the program SHARP²⁸. This map was greatly improved and showed us how to fit the YPWM motif. The YPWM fit was further verified by an anomalous-difference Fourier map calculated with data measured from selenomethionine (SeMet)-substituted protein; this map showed the positions of the three substituted seleniums, including the one in the YPWM motif. The refinement of the structure was extended to 2.4 Å resolution using the native 1 data, and the structure verified through extensive simulated annealing omit maps. Finally, 110 water molecules were added from the inspection of $F_o - F_c$ maps. The final refined structure has good stereochemistry, with the Ramachandran plot showing 84.5% of the residues in the core allowed regions and no residues in the disallowed or generously allowed regions.

Received 18 December 1998; accepted 3 February 1999.

- Lewis, E. B. A gene complex controlling segmentation in *Drosophila*. *Nature* **276**, 565–570 (1978).
- McGinnis, W. & Krumlauf, R. Homeobox genes and axial patterning. *Cell* **68**, 283–302 (1992).
- Mann, R. S. & Chan, S.-K. Extra specificity from extradenticle: the partnership between HOX and exd/pbx homeodomain proteins. *Trends Genet.* **12**, 258–262 (1996).
- Mann, R. S. The specificity of homeotic gene function. *Bioessays* **17**, 855–863 (1995).
- Gehring, W. J., Affolter, M. & Burglin, T. Homeodomain proteins. *Annu. Rev. Biochem.* **63**, 487–526 (1994).
- Wolberger, C. Homeodomain interactions. *Curr. Opin. Struct. Biol.* **6**, 62–68 (1996).
- Lu, Q. & Kamps, M. Structural determinants of Pbx1 mediating cooperative DNA-binding with pentapeptide-containing HOX proteins. *Mol. Cell. Biol.* **16**, 1632–1640 (1996).
- Chan, S.-K. & Mann, R. S. A structural model for an extradenticle-HOX-DNA complex accounts for the choice of HOX protein in the heterodimer. *Proc. Natl Acad. Sci. USA* **93**, 5223–5228 (1996).
- Chang, C. P., Brocchieri, L., Shen, W. F., Largman, C. & Clearly, M. L. Pbx modulation of Hox homeodomain amino-terminal arms establishes different DNA-binding specificities across the Hox locus. *Mol. Cell. Biol.* **16**, 1734–1745 (1996).
- Li, T., Stark, M. R., Johnson, A. D. & Wolberger, C. Crystal structure of the MATa1/MAT alpha 2 homeodomain heterodimer bound to DNA. *Science* **270**, 262–269 (1995).
- Chan, S. K., Jaffe, L., Capovilla, M., Botas, J. & Mann, R. S. The DNA binding specificity of Ultrabithorax is modulated by cooperative interactions with extradenticle, another homeoprotein. *Cell* **78**, 603–615 (1994).
- Fraenkel, E. & Pabo, C. O. Comparison of X-ray and NMR structures for the Antennapedia homeodomain-DNA complex. *Nature Struct. Biol.* **5**, 692–697 (1998).
- Izpisua-Belmonte, J. C., Falkenstein, H., Dolle, P., Renucci, A. & Duboule, D. Murine genes related to *teH Drosophila* AbdB homeotic gene are sequentially expressed during development of the posterior part of the body. *EMBO J.* **10**, 2279–2289 (1991).
- Johnson, F. B., Parker, E. & Krasnow, M. A. Extradenticle protein is a selective cofactor for the *Drosophila* homeotics: role of the homeodomain and YPWM amino acid motif in the interaction. *Proc. Natl Acad. Sci. USA* **92**, 739–743 (1995).
- Chan, S.-K., Pöppel, H., Krumlauf, R. & Mann, R. S. An extradenticle-induced conformational change in a HOX protein overcomes an inhibitory function of the conserved hexapeptide motif. *EMBO J.* **15**, 2477–2488 (1996).
- Rauskolb, C., Smith, K., Peifer, M. & Wieschaus, E. Extradenticle determines segmental identities throughout development. *Development* **121**, 3663–3671 (1995).
- Klemm, J. D. & Pabo, C. O. Oct-1 POU domain-DNA interactions: cooperative binding of isolated subdomains and effects of covalent linkage. *Genes Dev.* **10**, 27–36 (1996).
- Phelan, M. L. & Featherstone, M. S. Distinct HOX N-terminal arm residues are responsible for specificity of DNA recognition by HOX monomers and HOX-PBX heterodimers. *J. Biol. Chem.* **272**, 8635–8643 (1997).
- Aggarwal, A. K., Rodgers, D. W., Drott, M., Ptashne, M. & Harrison, S. C. Recognition of a DNA operator by the repressor of phage 434: a view at high resolution. *Science* **242**, 899–907 (1988).
- Otwinowski, Z. *et al.* Crystal structure of trp repressor/operator complex at atomic resolution. *Nature* **335**, 321–329 (1988).
- Tan, S. & Richmond, T. J. Crystal structure of the yeast MATalpha2/MCM1/DNA ternary complex. *Nature* **391**, 660–666 (1998).
- Janin, J., Miller, S. & Chotia, C. Surface, subunit interfaces and interior of oligomeric proteins. *J. Mol. Biol.* **204**, 155–164 (1988).
- Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
- Collaborative Computational Project, N. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
- Hirsch, J. A. & Aggarwal, A. K. Structure of the even-skipped homeodomain complexed to AT-rich DNA: new perspectives on homeodomain specificity. *EMBO J.* **14**, 6280–6291 (1995).
- Brunker, A. T. *XPLOR Version 3.1 Manual* (Yale Univ., New Haven, 1993).
- Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**, 110–119 (1991).
- de La Fortelle, E. & Bricogne, G. Maximum-likelihood heavy atom parameter refinement for the multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Methods Enzymol.* **276**, 472–494 (1997).

- Evans, S. V. Setor: hardware lighted three-dimensional solid model representations of macromolecules. *J. Mol. Graph.* **11**, 134–138 (1993).
- Nicholls, A., Sharp, K. A. & Honig, B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**, 281–296 (1991).

Acknowledgements. We thank the staff at CHESS for help with data collection; C. Escalante for advice on protein purification; L. Shapiro for help with map calculations; and T. Jessell and L. Shapiro for comments on this manuscript. J.M.P. thanks the Posen family for their hospitality during trips to CHESS. This work was supported by NIH grants to A.K.A. and R.S.M. R.S.M. is a Scholar of the Leukemia Society of America.

Correspondence and requests for material should be addressed to A.K.A. (e-mail: aggarwal@inka.mssm.edu). Coordinates have been deposited in the Brookhaven Protein Database (accession number 1B81).

erratum

Structural basis for activation of the titin kinase domain during myofibrillogenesis

Olga Mayans, Peter F. M. van der Ven, Matthias Wilm, Alexander Mues, Paul Young, Dieter O. Fürst, Matthias Wilmanns & Mathias Gautel

Nature **395**, 863–869 (1998)

D.O.F.'s full address should have included the Institut für Zoophysiology und Zellbiologie (University of Potsdam, Lennéstrasse 7a, 14471 Potsdam, Germany); lane 1 of Fig. 5a referred to transfected kin4 (not in4); and in Fig. 6b, the panels referred to as "top" and "bottom" should have been left and right panels, respectively. □

correction

Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*

Richard A. Alm, Lo-See L. Ling, Donald T. Moir, Benjamin L. King, Eric D. Brown, Peter C. Doig, Douglas R. Smith, Brian Noonan, Braydon C. Guild, Boudewijn L. deJonge, Gilles Carmel, Peter J. Tummino, Anthony Caruso, Maria Uria-Nickelsen, Debra M. Mills, Cameron Ives, Rene Gibson, David Merberg, Scott D. Mills, Qin Jiang, Diane E. Taylor, Gerald F. Vovis & Trevor J. Trust

Nature **397**, 176–180 (1999)

Typographical errors in Table 1 caused the transposition of some numbers between the *H. pylori* 26695 and J99 columns. The affected rows should read as follows.

<i>vacA</i> genotype:	26695, <i>sla/ml</i> ; J99, <i>slb/ml</i>
Functionally classified ORFs:	26695, 895; J99, 874
Conserved with no function ORFs:	26695, 290; J99, 275
<i>H. pylori</i> -specific ORFs:	26695, 367; J99, 346

In the table footnote, the coordinates of the second 26695 23S rRNA sequence should read 1,473,499–1,476,836. □

Combinatorial harvesters

New offerings from 15 manufacturers claim to make combinatorial chemistry easier, quicker, cleaner, cheaper — or perhaps unnecessary.

Heavy-duty shakers

From New Brunswick www.nbsc.com

Precision shakers designed with combinatorial chemistry in mind

Two new heavy duty 'combi-chem shakers' from New Brunswick Scientific are built to cope with high speeds, heavy workloads and sudden stops, with minimal wear or vibration. For use in an automated sampling station, the Innova Model 2130 can mix in a 3/4-inch stroke at up to 500 r.p.m., and then stop in exactly the same position each time. Once stopped, a robotics device can be used to fill or extract solutions from the reaction block. For high-speed mixings and washings, Model 2180 offers a shorter stroke and extended speed range up to 800 r.p.m. to give vigorous mixing and thorough washings.

Reader Service No. 100

ReactoStation

From Stem Corporation www.stemcorp.com

Combinatorial chemistry at high temperatures

Parallel synthesis reactions can be performed at up to 300 °C in the ReactoStation — previously 150 °C was regarded as a practical limit. Improved insulation keeps the stirring motors and electronics cool. Up to ten reactions can be heated and stirred within a very small footprint, saving space on the bench or in the fume cupboard. An optional module that fits on top of the station makes it possible to carry out reactions under reflux and near-inert conditions.

Reader Service No. 101



In the pink: ReactoStation can handle 300 °C.



Shake on it: Innova 2130 from New Brunswick.

384-well Microplates

From Packard www.packardinst.com

Assay miniaturization — 384 wells are ousting 96

Introduced last year, Packard Instrument Company's 384-well versions of the established 96-well OptiPlate, CulturPlate and ViewPlate systems provide new tools to miniaturize assays, conserve samples, reduce waste and increase throughput. The OptiPlate is an opaque (black or white), polystyrene microplate intended for isotopic, luminescence or fluorescence studies. CulturPlate is an opaque, sterile, tissue culture treated microplate designed for cell growth research. And ViewPlate is a black or white, clear-bottomed polystyrene microplate that allows microscopic visualization of the sample. These microplates are compatible with Packard's TopCount HTS and NXT microplate scintillation and luminescence systems, LumiCount microplate luminometer, FluoroCount microplate fluorescence reader and MultiPROBE II liquid handling system.

Reader Service No. 102

Flexys

From Genomic Solutions

www.genomicsolutions.com

A robotic workstation

The Flexys robotic workstation allows the user to rearrange a full library or to retrieve selected samples from the library and inoculate them into different positions in a new plate. The system utilizes a 24-pin pneumatic head and an Autoloader module and includes a lidlifter and optional refrigeration, both important for the librarian function. Flexys input and output holders accommodate eight 96- or 384-well microtitre plates.

Reader Service No. 103

LabChip

From Caliper

www.calipertech.com

Automated screening means big numbers, but small volumes

Caliper's Technology Access Program is aimed at drug companies needing to screen the enormous numbers of potential active compounds being generated by genomic studies and novel chemistry. The company's target is to have a system up and running during 1999 that is capable of performing more than 100,000 nanoscale assays per day.

Reader Service No. 104

ALPHAScreen

From Biosignal

www.biosignal.com

High-throughput screening using 384- or 1,536-well microplates

The ALPHAScreen platform gets its name from the assay technology involved — amplified luminescent proximity homogeneous assay. The system is aimed at the drug screening market where miniaturization is a key factor. It combines biology, liquid handling robotics and detection instrumentation to cope with ultra-high-throughput screening assays in 1,536-well microplates.

Reader Service No. 105

SMS

From Perkin-Elmer

www.perkin-elmer.com

Monitoring solid-phase synthesis using the infrared

Manufacturers Perkin-Elmer introduced their Synthesis Monitoring System (SMS) as a stand-alone analyser designed specifically for the pharmaceutical drug discovery market. Based on Fourier transform-infrared (FT-IR) technology, the SMS analyser verifies reaction sequence products before starting complex synthesis procedures and also confirms optimal synthesis conditions by estimating a reaction's



SMS: monitor synthesis on the monitor.

Nothing But Net.

Explore the boundaries of the web with *Nature* Web Matters. Short essays on topics of importance to web users written specifically for the scientific, web-using audience.

Essays will cover topics directly related to the web, as well as discuss technologies such as preprint servers and online databases. Examine the issues that will affect the electronic dissemination of scientific knowledge in the months and years to come.

Think **nature** First
www.nature.com

Finders Keepers.

Expanding your search for more information has never been easier. *Nature's* current and past issues are fully searchable online. References in the bibliography of the printed issue document are linked to abstracting services, including ISI's Web of Science and Medline. Or, let *Nature* online search for you - a list of related papers published in *Nature* and elsewhere are only a click away. You never have to leave our website to expand your mind.

Think **nature** First
www.nature.com

Chemiluminescent Substrates



A Super Spectrum of Possibilities

Pierce SuperSignal® Chemiluminescent Substrates for all your protein research needs.

Chemiluminescence

Western Blotting Applications

SuperSignal® West Pico Chemiluminescent Substrate*
Picogram-level sensitivity at an amazingly economical price
Product # 34080EL, 500 ml
(Formerly SuperSignal® Chemiluminescent Substrate for Western Blotting)

SuperSignal® West Dura Extended Duration Substrate*
24-hour signal duration and super sensitivity
Product # 34075EL, 100 ml
(Formerly SuperSignal® ULTRA Chemiluminescent Substrate)

SuperSignal® West Femto Maximum Sensitivity Substrate*
The most intense signal and the greatest sensitivity attainable with a chemiluminescent substrate
Product # 34095EL, 100 ml
(Formerly SuperSignal® BLAZE™ Chemiluminescent Substrate)

ELISA

SuperSignal® ELISA Pico Chemiluminescent Substrate*
Picogram-level detection for plate-based assays
Product # 37370EL, 100 ml
(Formerly SuperSignal® LBA for Luminometer-Based Applications)

SuperSignal® ELISA Femto Maximum Sensitivity Substrate*
A more sensitive substrate with enhanced low-end linearity
Product # 37375EL, 100 ml
(Formerly PowerSignal™ ELISA Chemiluminescent Substrate)
*Power Reading

Call for your **FREE** Substrates Selection Guide Poster!

Call 800-874-3723 for product information or 815-968-0747 for the name of your local distributor.



a Henry L. Pierce Company

3747 N. Meridian Rd. • PO Box 117 • Rockford, IL 61105 U.S.A. • Fax (815) 968-7216
E-mail: TD@pierce.com • Internet: <http://www.pierce.com> © Pierce Chemical Company, 1998

RECOMBINANT SERVICES



ProteinGenesys
Proteins for Sciences of Life

We Can Deliver Purified, Native & Glycosylated, You Just Name the Protein, We Do The Rest

Proteins for Research, Applied Protein Chemistry, Drug Discovery, Genomics, Immunochemistry and More
www.proteingenesis.com toll: +41 61/811 3357, 2803 fax

READER ENQUIRY NO. 5

Detect residual DNA with Wako's DNA Extractor Kit
or Isolate Genomic DNA with DNA Extractor WB Kit
Both use the Nal Method
No Phenol No Chloroform

Wako

Wako BioProducts
Wako Chemicals USA, Inc.
1600 Bellwood Road
Richmond, VA 23237
Telephone: (804) 271-7677
Facsimile: (804) 271-7791
(800) 992-9256
bioproducts@wakousa.com

READER ENQUIRY NO. 36599 Macmillan Magazines Ltd

READER ENQUIRY NO. 391

percentage completion. The software automatically compares spectra from 'before' and 'after' reaction stages. Later, users select functional group information, and highlight the spectral differences for correlation with the changes expected from the functional group selection.

Reader Service No. 106

Combi-Clamp

From Whatman, Inc. www.whatman.com

Combinatorial synthesis in a closed system

Combi-Clamp is designed for use with the Whatman 96-well 2-ml Unifilter filter bottom microplate. It allows the use of organic and other highly volatile solvent systems by eliminating problems associated with evaporation. This allows optimization of reaction times and conditions that otherwise would be difficult to achieve without a closed system. Combi-Clamp consists of a two-piece chemically resistant anodized aluminium block, PTFE sealing gaskets, and other associated hardware necessary to seal the 'reaction' within each well. When the syntheses are complete, the clamp is removed, and the filter bottom microplate

evacuated via a vacuum manifold, or otherwise processed as directed in the synthetic protocol.

Reader Service No. 107

FLIPR384

From Molecular Devices www.moldev.com

High-throughput fluorescence analysis

The FLIPR system (for Fluorimetric Imaging Plate Reader) performs fluorimetric analysis, such as the measurement of changes in intracellular calcium, on cells in a rapid-throughput real-time format.



FLIPR384: fluorescence determinations in bulk.

FLIPR384 is the latest application of this technology. Fluorescent light is captured from all wells of either a 384- or 96-well plate simultaneously, while a pipettor delivers compound to each well from up the three separate compound plates. High throughput is claimed, over 50,000 kinetic determinations in an 8-hour day using the 384-well format.

Reader Service No. 108

Pharmaceutical data

From Derwent Information Ltd

www.derwent.com

Information — free for 30 days

The new pharmaceutical information service, Derwent Discovery, now includes access to the Derwent Drug File, a journal-based biomedical database with entries going back to 1964. The second major information source included in the service is the Derwent World Drugs Alert. Derwent are offering 30 days' free access to these databases. Go to www.derwent-discovery.com, or e-mail to salesup@derwent.co.uk

Reader Service No. 109

MultiReactor

From RoboSynthon www.robosynthon.com

Two dozen reactions at once

The MultiReactor from Robosynthon, Inc., is a new combinatorial chemistry workstation for parallel synthesis. It is designed to run 24 synthetic reactions simultaneously, with vigorous magnetic stirring. Features include computer-controlled mixing and heating, viscous liquid and solid-phase programs. MultiReactor can be controlled by a computer RS-232 port for semi-automated control of mixing and temperature. Solid phase mixing is achieved with short stirring pulses. Viscous mixing is done with a slow increasing ramp. Mixing and heating programs can be done simultaneously.

Reader Service No. 110

Molecular-weight confirmation

From Hewlett-Packard www.hp.com

Liquid chromatograph/mass spectrometer hardware and software

The HP molecular-weight-confirmation system provides rapid molecular-weight



Source of information: MWs from HP.

ADVERTISEMENT

Turn Blu at the Possibilities

Pierce QuantaBlu™
Chemifluorescent Substrates

Fluorogenic HRP substrates for plate-based applications

QuantaBlu™ Fluorogenic Peroxidase Substrate
Ideal for use with peroxidase enzymes
Product # 15169EL, 250 ml

QuantaBlu™ NS/K
Specifically designed for nonstopped and kinetic assays
Product # 15162EL, 250 ml

Features/Benefits:

- More sensitive than TMB, OPD or ABTS Substrates
- Flexible – stopped, nonstopped or kinetic assays possible
- Large dynamic range (4 log peroxidase concentration range)
- Excellent 24-hour working solution stability
- Large Stokes shift; excitation/emission maxima of 325/420; range of 315-345/370-460
- Does not photobleach

Call for your FREE Substrates Selection Guide Poster!
Call 800-374-3723 for product information or
815-968-9747 for the name of your local distributor.

PIERCE
a PerkinElmer Company

2747 N. Meridian Rd. • PO Box 117 • Rockford, IL 61106 U.S.A. • Fax: 815-968-1310
E-mail: TA@piercelabs.com • Internet: <http://www.piercelabs.com> © Pierce Chemical Company, 1998

READER ENQUIRY NO. 366

confirmation of combinatorial libraries, taking the process through from sample analysis to reporting. At the core of the system are the HP 1000 HPLC, HP ChemStation, and the HP 1100 LC/MSD, with the necessary software provided. HP claim that using this combination even chemists without mass-spectrometry expertise can soon achieve fast confirmation of combinatorial methods and efficient verification of large combinatorial-synthesis projects.

Reader Service No. 111

Titan microtitre plates

From Radleys www.radleys.co.uk

PTFE in microplate form for stability at high temperatures

Titan is a range of 96- and 24-well microtitre plates and lids made from pure PTFE, recommended for combinatorial chemistry, compound libraries and all micro reactions. PTFE offers thermal stability at up to 300 °C, and does not become brittle in liquid helium. Titan PTFE plates can be used for storage of compound libraries, collection of aggressive chemicals or compounds and for heated or cooled

micro reactions in combinatorial chemistry.

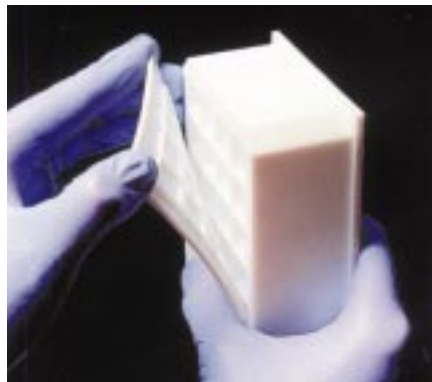
Reader Service No. 112

Model 496

From Advanced ChemTech www.peptide.com

Synthesis in 8 to 96 wells

ACT's Model 496 multiple organic synthesizer (MOS) is a fully automated dual-arm robotics synthesizer which can make up to 96 different compounds in a single run. Reaction blocks with 40, 16 and 8 wells are also available. The individual



Hot stuff: PTFE features in Radleys' microplates.

working volumes increase as the number of wells is reduced. The Model 496 MOS has an inert gas atmosphere and a reaction heating and cooling system. The temperature range of the unit is -70° to +150° C.

Reader Service No. 113

Contract research

From MDS Panlabs www.panlabs.com

Take out a contract

MDS Panlabs is a contract research organization providing scientific research services supporting the pharmaceuticals industry. Specific services include: pharmacological profiling; strain and yield improvement using molecular or classical approaches; biosafety assessment of biopharmaceuticals; bioanalytical measurements of drug metabolism; bioavailability and pharmacokinetics; and compound synthesis, isolation and analysis.

Reader Service No. 114

These notes are compiled in the Nature office from information provided by the manufacturers. For more details, fill in the reader service card bound inside the journal.

ADVERTISEMENTS

Pierce 1-Step™ Colorimetric Substrates

Ready-to-use, pre-formulated colorimetric substrates.

1-Step™ Substrates are available in the following formulations:

Slow TMB Product # 34024EL, 250 ml Turbo TMB Product # 34022EL, 250 ml ABTS Product # 37615EL, 250 ml Chloronaphthol Product # 34012EL, 250 ml	NBT/BCIP Product # 34042EL, 250 ml PNPP Product # 37621EL, 100 ml Features/Benefits: • One component • Ready-to-use • Stable
---	---

Call for your FREE Substrates Selection Guide Poster!
Call 800-874-3723 for product information or
815-966-0747 for the name of your local distributor.

PIERCE
a PerkinElmer Life Sciences Company

3747 N. Meridian Rd. • PO Box 117 • Rockford, IL 61125 U.S.A. • Fax: 815-966-7316
E-mail: 140@piercelife.com • Internet: <http://www.piercelife.com> © Pierce Chemical Company, 1998

"Beside good science, there is much humour and much of interest outside the field of science. ... In my case, it kept me awake and delighted until the small hours"

Keith J. Laidler, Professor Emeritus of Chemistry, Ottawa University

The Bedside Book Genius and Eccentricity in Science 1869–1953

Edited by Walter Gratzer, with a foreword by Stephen Jay Gould

A fascinating selection of articles from Nature, offering a panorama of science seen against the background of 19th and 20th century history.



Price: £19.95/US\$29.95 + postage
Discount rate available for current subscribers.
To order, or for more information, contact: Nature Marketing
Tel. +44 (0) 171 843 4985
Fax. +44 (0) 171 843 4998
E-mail: subscriptions@nature.com

READER ENQUIRY NO. 367